

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Desenvolvimento de um Sistema de Apoio à Decisão para a manutenção preditiva dos ativos de uma subestação elétrica

Afonso Neves Caldas



Mestrado Integrado em Engenharia Informática e Computação

Orientador: João Mendes Moreira (PhD)

Julho de 2015

© Afonso Neves Caldas, 2015

Desenvolvimento de um Sistema de Apoio à Decisão para a manutenção preditiva dos ativos de uma subestação elétrica

Afonso Neves Caldas

Mestrado Integrado em Engenharia Informática e Computação

Aprovado em provas públicas pelo Júri:

Presidente: João Pascoal Faria (PhD)

Vogal Externo: José Moreira (PhD)

Orientador: João Moreira (PhD)

21 de julho de 2015

Resumo

Numa empresa como a EDP Distribuição, a ocorrência de falhas nos ativos de uma subestação pode ter consequências graves. Os equipamentos próximos podem também ser afetados e podem até causar cortes de energia afetando os consumidores.

A EDP Distribuição desenvolveu algoritmos para cálculo de estimativas: (1) do Índice de Vida, (2) da Probabilidade de Falha e (3) do tempo de vida restante dos principais ativos de uma subestação. Estas métricas podem dar uma perspectiva clara e facilmente compreensível do estado do ativo e ajudar no processo de manutenção de uma rede de Subestações.

Os algoritmos têm vários parâmetros provenientes de diferentes fontes de informação como testes de diagnóstico, inspeções de rotina ou monitorização *online*, o que faz com que seja difícil aplica-los a todas as Subestações. No trabalho desenvolvido criou-se uma plataforma protótipo onde, de modo simples, compilando toda a informação disponível, se cria um Sistema de Apoio à Decisão que calcula os indicadores referidos e apresenta-os de uma forma prática e intuitiva. Através desta aplicação, um operador tem acesso a uma rede de Subestações e pode facilmente verificar a condição dos ativos de cada uma das Subestações.

São também estudadas diferentes aplicações potenciais de algoritmos de *Data Mining* para detetar padrões nos dados existentes e providenciar informação valiosa para o planeamento de manutenção. Para o caso específico da estimação da Probabilidade de Falha foram testados diferentes métodos que permitiram melhorar as previsões dadas pelo algoritmo existente na empresa. No caso de estudo abordado, os métodos de *Survival Analysis* aplicados foram capazes de fazer uma previsão da probabilidade de um disjuntor falhar consoante a sua idade, tendo obtido resultados significativamente melhores que o algoritmo empírico num conjunto de disjuntores com tempo de vida conhecido.

O objetivo final do trabalho desenvolvido é auxiliar a transição de um modelo de manutenção baseado no tempo para um modelo de manutenção preditivo das Subestações, fornecendo a informação necessária para aplicar medidas preventivas antes que uma falha ocorra.

Abstract

In a company like EDP Distribuição, the occurrence of faults in a Substation's assets can have dire consequences. It may cause damages to the surrounding equipment and ultimately cause power outages affecting the consumers.

EDP Distribuição has developed algorithms to calculate an estimation of: (1) the Health Index, (2) the Failure Probability and (3) the remaining lifetime of a Substation's assets. These metrics give a clear and understandable picture of the asset's state and help in the management of a Substation network.

The algorithms have various parameters from different sources like diagnostic tests, routine inspections or online monitoring, which makes it a difficult task to apply them to all the Substations. A prototype of a Decision Support System was implemented, it uses the necessary inputs to calculate the referred metrics and present them in an understandable and practical way. Using this application, an operator has access to a subset of the network of Substations and can easily check the condition of each Substation's assets.

Different potential applications of Data Mining algorithms are also studied to detect patterns in the existing data and provide valuable inputs to the maintenance planning. To the specific case of estimating the Failure Probability various methods were tested and they were able to improve the predictions made by the algorithm created in the company. In the designed case study, the Survival Analysis applied were able to make a prediction of the probability of failure of a Circuit Breaker, obtaining significantly better results than the empirical algorithms in a set of Circuit Breakers with known lifetime.

The final goal is to allow a move from time-based maintenance to a predictive maintenance of the Substations, giving the information needed to apply preventive measures before a fault occurs.

Agradecimentos

Quero agradecer à minha família, aos meus pais e irmã, pelo apoio incondicional e ânimo dado ao longo do meu percurso académico.

Um agradecimento ao meu orientador Prof. João Mendes Moreira pela disponibilidade e pelos conselhos dados na realização desta dissertação.

Queria, também, agradecer a todos os colaboradores da EDP Distribuição, em especial ao Eng.º João Vasco Ferreira, Eng.º Miguel Freitas e Eng.º Rúben Neves pelo apoio prestado, pela receção e pelas condições oferecida durante a minha estadia na EDP.

Por fim, não podia deixar de agradecer a todos os meus amigos pela sua amizade e encorajamento.

Afonso Neves Caldas

“The best way to predict the future is to invent it.”

Alan Kay

Conteúdo

Introdução.....	2
1.1 Contexto	2
1.2 Motivação e Objetivos	3
1.3 Estrutura da Dissertação.....	4
Conceitos Relevantes.....	5
2.1 Gestão de Ativos	5
2.1.1 Manutenção Corretiva	6
2.1.2 Manutenção Preventiva	6
2.1.3 Manutenção Preditiva.....	6
2.2 Data Mining	13
2.2.1 Classificação e Regressão	13
2.3 Survival Analysis	15
2.3.1 Função de Sobrevida.....	15
2.3.2 Função de Risco	16
2.3.3 Kaplan-Meier	16
2.3.4 Cox Proportional Hazards	17
2.3.5 Métodos Paramétricos	18
2.3.6 Árvore de Decisão.....	21
2.3.7 Random Survival Forest.....	21
Gestão de Ativos numa subestação elétrica	23
3.1 Principais Ativos de uma subestação	25
3.1.1 Transformador de Potência	25
3.1.2 Disjuntor.....	26
3.1.3 Sistemas de Proteção, Comando e Controlo	28
3.1.4 Sistemas de Alimentação	30
3.2 Algoritmos desenvolvidos na EDP	31
3.2.1 Índice de Saúde	31
3.2.2 Probabilidade de Falha	35

3.2.3	Tempo de Vida Restante	39
3.2.4	Risco.....	39
3.3	Análise de ensaios em disjuntores.....	41
Aplicações de Data Mining para Manutenção Preditiva de uma subestação		45
4.1	Revisão do Estado da Arte	46
4.1.1	Índice de Saúde	46
4.1.2	Probabilidade de Falha	52
4.1.3	Tempo de vida restante	54
4.1.4	Conclusões	56
4.2	Estimação da Probabilidade de Falha de um disjuntor através de <i>Survival Analysis</i>	57
4.2.1	Preparação dos dados	58
4.2.2	Metodologia adotada	60
4.2.3	Resultados	61
Implementação do protótipo do Sistema de Apoio à Decisão.....		65
5.1	Especificação de Requisitos	65
5.2	Tecnologias utilizadas	66
5.3	Modelo de dados	66
5.4	Interfaces gráficas	69
5.4.1	Portal	69
5.4.2	Painel de Administração.....	75
Conclusões e Trabalho futuro		81
6.1	Conclusões	81
6.2	Desenvolvimentos futuros.....	82
Referências.....		84

Lista de Figuras

Figura 1 - Matriz de Risco	12
Figura 2 - Transformador de Potência	26
Figura 3 - disjuntor	27
Figura 4 - Cela de Média Tensão	27
Figura 5 - Intelligent Equipment Device	28
Figura 6 - Unidade Central de Processamento	29
Figura 7 - Baterias de Tensão Contínua	30
Figura 8 - Estrutura do algoritmo	31
Figura 9 - Função da probabilidade de falha de um disjuntor consoante a idade	37
Figura 10 - Matriz de Risco EDP Distribuição	40
Figura 11 - Aplicabilidade dos indicadores de condição	40
Figura 12 - Número de ensaios por meio de interrupção de disjuntor	42
Figura 13 - Resistência de Contacto dos polos por meio de interrupção	43
Figura 14 - Resistência de Contacto de disjuntores a Óleo e de Média Tensão	44
Figura 15 - Modelo difuso para criticidade térmica (Arshad, Islam, and Khaliq 2014)	47
Figura 16 - Aplicação de Redes Neurais para avaliação de transformadores (Mukti, Pamuji, and Munir 2014)	48
Figura 17 - Rede Neuronal proposta por Abu-Elanien et al (2001)	49
Figura 18 - <i>Multi-Layer</i> SVM para deteção de falha (Bacha, Souahlia, and Gossa 2012)	53
Figura 19 - Contribuição da idade do ativo nos algoritmos existentes	57
Figura 20 - Meio de interrupção dos disjuntores abatidos	58
Figura 21 - Meio de interrupção dos disjuntores em serviço	59
Figura 22 - Relação entre o IBS de cada modelo e o número de indivíduos censurados	62
Figura 23 - Relação entre o IBS de cada modelo e o número de indivíduos não censurados	64
Figura 24 - Diagrama de classes para modelação dos algoritmos	67
Figura 25 - Diagrama de classes da rede de ativos	68
Figura 26 - Página Inicial	70

Figura 27 - Pesquisa de subestação por nome ou por SAP ID	70
Figura 28 - Página Inicial - Ativos em risco	71
Figura 29 - Pesquisa em árvore dos ativos	71
Figura 30 - Página de subestação	72
Figura 31 - Página de Ativo - Indicadores principais	73
Figura 32 - Página de Ativo - Vista detalhada dos indicadores	74
Figura 33 - Página de Ativo - Histórico	74
Figura 34 - Página Inicial do Painel de Administração	75
Figura 35 - Página de Tipo de Ativo	76
Figura 36 - Página de Tipo de Ativo - Definição dos componentes	77
Figura 37 - Página de Tipo de Ativo - Definição de Parâmetro	77
Figura 38 - Página de criação de subestação	78
Figura 39 - Página de criação de Ativo - Seleção de Parâmetro	79
Figura 40 - Página de criação de Ativo	79
Figura 41 - Página de criação de Ativo - Atribuição de valor	79
Figura 42 - Página de criação de funções	80

Lista de Tabelas

Tabela 1 - Critérios para formulação do Índice de Saúde (Hjartarson and Otal 2006)	11
Tabela 2 - Matriz de Confusão	14
Tabela 3 - Modos de falha de um disjuntor	33
Tabela 4 - Parâmetros impactantes no Índice de Saúde da Câmara de corte	34
Tabela 5 - Parâmetros impactantes no Índice de Saúde do Comando	34
Tabela 6 - Fatores impactantes nos modos de falha de um disjuntor	36
Tabela 7 - Peso atribuído aos fatores impactantes nos modos de falha de um disjuntor	36
Tabela 8 - Relação entre a probabilidade de falha de um disjuntor e a sua idade	37
Tabela 9 - Fatores externos impactantes num disjuntor	38
Tabela 10 - Tempo de vida máxima por tecnologia de disjuntor	39
Tabela 11 - Criticidades de um Transformador (Arshad, Islam, and Khaliq 2014)	47
Tabela 12 - Comparação de SVM com e sem análise de correlação (Ashkezari et al. 2014)	51
Tabela 13 - Comparação de modelos SVM com diferente número de variáveis (Ashkezari et al. 2014)	51
Tabela 14 - Código dos tipos de falha por DGA (Bacha, Souahlia, and Gossa 2012)	53
Tabela 15 - Comparação de métodos de IA para detecção de falhas (Bacha, Souahlia, and Gossa 2012)	54
Tabela 16 - Dados de sobrevivência de disjuntores	59
Tabela 17 - Erro dos modelos por número de indivíduos censurados	61
Tabela 18 - <i>Brier Skill Score</i> dos modelos com 5000 indivíduos censurados	62
Tabela 19 - Erro dos modelos por número de indivíduos não censurados	63

Abreviaturas e Símbolos

BSS	Brier Skill Score
CIGRE	Conseil International des Grands Réseaux Électriques
EDP	Energias de Portugal
IBS	Integrated Brier Score
IEC	International Electrotechnical Commission
IS	Índice de Saúde
PdF	Vida Restante
SF6	Hexafluoreto de enxofre
SVM	Support Vector Machine
VR	Vida restante

Capítulo 1

Introdução

As empresas de distribuição de energia elétrica têm a missão crítica de garantir o fornecimento de energia aos consumidores finais. A interrupção do serviço pode ter graves consequências na economia global, no entanto, as distribuidoras de energia são empresas que desenvolvem o seu trabalho dentro de limitações tecnológicas e financeiras e que procuram manter um equilíbrio económico sustentável. É, portanto, fundamental assegurar a fiabilidade do fornecimento de energia minimizando os custos associados.

No presente capítulo é feito o enquadramento do tema desta dissertação de mestrado e são detalhadas as motivações e os objetivos do trabalho desenvolvido. Por fim, é explicitada a estrutura deste documento.

1.1 Contexto

O trabalho foi desenvolvido no âmbito da EDP Distribuição que é a empresa responsável pela operação da Rede de Distribuição de Energia elétrica em Média Tensão e Alta Tensão de Portugal Continental. A sua função é a distribuição de energia elétrica aos seus clientes. A Rede de Distribuição é composta por diversas infraestruturas e equipamentos que têm que ser geridos de forma a garantir a qualidade do serviço prestado pela empresa.

O Departamento de Manutenção e Controlo da EDP Distribuição é responsável por definir estratégias de gestão dos ativos presentes na Rede de Distribuição. A sua função é fundamental para a empresa atinja níveis de eficiência elevados em termos funcionais e económicos. Se registarmos que muitos dos seus equipamentos, nomeadamente transformadores, têm uma média de idade superior a 30 anos, facilmente percebemos a importância deste departamento e dos métodos de gestão utilizados no controlo da vida útil dos ativos.

O tema proposto foca-se nas Subestações elétricas AT/MT e nos seus equipamentos, que têm como função transformar a tensão da corrente elétrica de Alta para Média tensão. Para responder às necessidades da empresa e assegurar um modelo de gestão mais eficiente para o elevado número de ativos afetos à operação, foi criado o projeto “subestação 5M”. Este projeto visa uma maior eficiência da operação através de Subestações: mais simples, mais pequenas, mais económicas, mais fáceis de manter e mais rápidas de construir.

No âmbito desta iniciativa, foram desenvolvidos modelos para a caracterização da condição dos principais ativos de uma subestação elétrica através de algoritmos que estimam os seguintes indicadores:

- Índice de Saúde;
- Probabilidade de Falha;
- Tempo de vida restante.

Através destes indicadores pretende-se analisar os ativos de uma forma objetiva, permitindo dar prioridade às intervenções de manutenção e a otimização da gestão das subestações.

1.2 Motivação e Objetivos

Uma falha na Rede de Distribuição pode ter impacto não só na EDP Distribuição mas também em vários quadrantes da sociedade. A interrupção do fornecimento de energia pode provocar grandes despesas aos consumidores e, no pior caso, causar o ferimento ou morte de pessoas. Tudo isto afeta gravemente a reputação da operadora responsável e resulta em grandes prejuízos. Assim sendo, estas empresas são intimamente dependentes dos seus ativos.

A EDP Distribuição tem, portanto, uma gestão focada nos seus ativos, garantindo a segurança e a qualidade do serviço, não descurando o seu equilíbrio financeiro. Para que a empresa atinja os seus objetivos é fundamental rentabilizar ao máximo esses ativos, criando modelos de gestão e operação que evitem falhas que comprometam a Rede de Distribuição, que aumentem a produtividade e, simultaneamente, reduzam os custos.

Neste contexto, são aplicados processos de manutenção preditiva de forma a atuar antes que ocorra uma falha e a estender o tempo de vida útil dos ativos, em contraponto a um tipo de manutenção corretiva que traria mais custos para a empresa e prejuízos para os seus clientes.

O objetivo principal desta dissertação é a implementação de um protótipo de um Sistema de Apoio à Decisão que permita a um operador aceder facilmente às subestações da Rede de Distribuição nacional e verificar a condição dos seus ativos. A aplicação automatiza os algoritmos desenvolvidos pela EDP Distribuição para estimação do Índice de Saúde, da Probabilidade de Falha e do tempo de vida restante. Este protótipo servirá de prova de conceito, auxiliando a especificação de um sistema informático a ser aplicado a nível industrial e a avaliar a sua utilidade, recursos necessários e aplicabilidade.

Outro objetivo é a revisão de abordagens na área de *Data Mining* que possam complementar os algoritmos existentes a ser implementados no contexto da EDP Distribuição.

A metodologia adotada para a realização dos objetivos propostos foi a seguinte:

- Identificação e reconhecimento da estrutura, do funcionamento e das características dos ativos de uma subestação elétrica através de visitas aos locais e de revisão da literatura especializada;
- Análise do modelo de gestão aplicado na manutenção das Subestações;
- Estudo e caracterização dos algoritmos aplicados e da natureza dos dados utilizados;
- Revisão do Estado da Arte de aplicações de *Data Mining* para a manutenção preditiva de Subestações;
- Desenvolvimento de caso de estudo para avaliar a aplicação das abordagens identificadas;
- Elaboração e desenvolvimento do protótipo;
- Análise das potencialidades e da aplicabilidade do protótipo.

1.3 Estrutura da Dissertação

Este documento inclui, para além da introdução que é o primeiro capítulo, mais cinco capítulos. No segundo capítulo são apresentados alguns conceitos fundamentais relativos à gestão de ativos, ao *Data Mining* e à *Survival Analysis*.

No terceiro capítulo, é aprofundado o problema da gestão dos ativos de uma subestação elétrica, sendo identificados os ativos em questão e os algoritmos desenvolvidos na EDP e respetivos indicadores. Em particular, é desenvolvido o conceito de manutenção preditiva. É também discutida uma análise estatística feita com o intuito de melhorar a caracterização de um disjuntor.

No capítulo seguinte é revisto o Estado da Arte na área de *Data Mining* sobre a estimação dos indicadores referidos nos principais ativos de uma subestação elétrica. É feita uma análise da aplicabilidade das abordagens identificadas e é apresentado um caso de estudo relativo à estimação da probabilidade de falha de um disjuntor. É explicitada sua integração no algoritmo, os dados utilizados, a metodologia de avaliação e os resultados obtidos.

No quinto capítulo é relatado todo o processo de desenvolvimento e implementação da aplicação, especificando-se os requisitos, as tecnologias utilizadas, modelo de dados e interfaces gráficas implementadas.

Por último, no sexto capítulo, são apresentadas as principais conclusões e é feita uma análise crítica ao protótipo implementado assim como ao restante trabalho desenvolvido. Identificam-se ainda linhas futuras de desenvolvimento com vista à melhoria do protótipo no âmbito de um Sistema de Apoio à Decisão num modelo de Manutenção Preditiva.

Capítulo 2

Conceitos Relevantes

Neste capítulo são desenvolvidos alguns conceitos relevantes para o trabalho realizado. A primeira secção é sobre Gestão de Ativos, apresentando os vários tipos de manutenção e explicando o propósito dos indicadores Índice de Saúde, Probabilidade de Falha, Tempo de Vida Restante e Risco. Na secção seguinte é introduzido o *Data Mining* e alguns problemas relevantes. A última secção é sobre *Survival Analysis*, as características do problema que esta área estatística pretende resolver bem como alguns métodos utilizados para esse efeito.

2.1 Gestão de Ativos

Neste contexto, um ativo pode ser definido como um recurso que cumpre uma função ou fornece um serviço. Para empresas e indústrias cujo sucesso é baseado nos seus ativos é essencial que a sua gestão seja planeada com vista a rentabilizá-los o máximo possível. Isto envolve definir uma estratégia para operar, preservar, melhorar e expandir os ativos ao longo do seu ciclo de vida. No caso dos ativos físicos, este processo tem várias fases: adquirir o ativo, a sua instalação, operação, manutenção, remodelação e o seu retiro (Anders, Otal, and Hjartarson 2006).

Para que uma gestão correta e criteriosa seja possível é necessário analisar de forma detalhada os ativos, as suas propriedades e dependências. Uma grande quantidade de informação tem que ser recolhida, analisada e interpretada para que se possa adotar medidas de gestão para responder às necessidades, problemas e oportunidades que surjam.

Um dos grandes alicerces da gestão de ativos é a sua manutenção. Através deste processo, pretende-se aumentar ao máximo o tempo de vida de um ativo, assegurando que este continue a cumprir as suas funções nas melhores condições possíveis e no máximo de tempo possível. Para isto devem ser tomadas medidas para prevenir a ocorrência de falhas no ativo ou para o restaurar.

No entanto, para que a gestão do ativo tenha em conta os objetivos transversais da empresa, tem que ser feito o balanço entre os benefícios de uma medida de manutenção recomendada para um ativo e os custos dessa medida.

De seguida serão apresentadas três estratégias de manutenção: Manutenção Corretiva, Manutenção Preventiva e Manutenção Preditiva.

2.1.1 Manutenção Corretiva

A Manutenção Corretiva é efetuada depois de um ativo falhar. Quando o ativo falha pode causar interrupções no serviço e afetar a *performance* do sistema, sendo necessário reparar o ativo para que este possa voltar a cumprir a sua função.

Este tipo de manutenção pode ser planeado ou não planeado, dependendo do método de gestão e das consequências de o equipamento falhar. Caso este seja o único método de manutenção então é uma Manutenção Reativa, que não exige planeamento e tem custos iniciais menores mas pode gerar grandes prejuízos. É difícil orçamentar os gastos devido à imprevisibilidade das falhas e dos possíveis custos de reparação. Além disso, ao apenas intervir quando ocorre uma falha está-se a incorrer nos riscos associados à ocorrência da falha, como danos colaterais em outros ativos. Como tal, este tipo de manutenção é mais recomendado para ativos com uma função pouco crítica e que sejam facilmente substituídos ou reparados.

2.1.2 Manutenção Preventiva

As técnicas de Manutenção Preventiva são utilizadas para descobrir deficiências e providenciar as condições necessários a um ativo para que esta possa continuar a cumprir a sua função (Gulati and Smith 2009). Têm o objetivo de reduzir a probabilidade de um ativo falhar e de evitar a sua degradação. Pode ser necessário substituir uma peça ou subcomponente do ativo segundo um critério pré-determinado, por exemplo mudar o óleo de um carro quando este atinge uma certa quilometragem.

Estas técnicas são aplicadas de forma sistemática, não tendo em conta o estado do ativo ou o seu rendimento, e requerem um planeamento inicial e uma quantidade constante de trabalho e materiais. São ideais para ativos com uma função crítica, em que todas as falhas possam ser prevenidas com manutenção constante e cuja probabilidade aumente com o tempo ou com o uso do ativo. É importante analisar a eficácia dos critérios definidos aplicados para que a manutenção não seja excessiva, interrompendo desnecessariamente o funcionamento do ativo.

2.1.3 Manutenção Preditiva

Tal como a Manutenção Preventiva, a Manutenção Preditiva tem em vista reduzir a probabilidade de falha de um ativo. A diferença é que a Manutenção Preditiva se baseia no estado

do ativo em vez de condições pré-determinadas. O objetivo é prever o momento em que um ativo vai falhar e, também, em atuar para prevenir essa falha. Para que tal seja possível, é fundamental monitorizar o ativo e compreender de forma profunda o seu funcionamento e a evolução do seu estado para prever a falha com eficácia e com tempo suficiente para atuar.

Dependendo das características do ativo, pode ser bastante complexo aplicar este tipo de manutenção mas, caso se consiga fazê-lo com sucesso, esta estratégia pode ser a que se traduz em menos gastos. Por um lado, como as falhas passam a ocorrer com menos frequência, são utilizados menos recursos e menos horas para reparar o ativo e este fica menos tempo fora de serviço. Por outro lado, realizando as intervenções de manutenção apenas quando o estado de ativo o justificar, reduz-se o número de intervenções e, consequentemente, os custos de manutenção.

O protótipo criado pretende dar apoio a este tipo de manutenção preditiva, pelo que se desenvolve mais este conceito.

Na manutenção preditiva é necessário considerar se as falhas de um ativo podem efetivamente ser previstas e se os custos associados com a monitorização do ativo justificam os ganhos em relação à Manutenção Preventiva ou à Corretiva.

Os passos tipicamente necessários para a Manutenção Preditiva de um ativo são (Insights 2003):

- Identificar que amostras e propriedades do ativo se deve monitorizar;
- Recolher as amostras especificadas continuamente ou periodicamente;
- Analisar medidas para detetar possíveis degradações no ativo e planear a sua manutenção personalizada;
- Verificar se o ativo cumpre os *standards* pré-definidos para continuar em funcionamento;
- Aplicar medida corretiva para melhorar a condição do ativo caso seja necessário.

Alguns indicadores relevantes para este tipo de manutenção serão agora apresentados.

2.1.3.1 Índice de Saúde

O Índice de Saúde é uma métrica utilizada na gestão de ativos para refletir a condição e o estado de degradação e de envelhecimento de um ativo (Dominelli 2004; Hjartarson and Otal 2006). É um valor facilmente compreensível, enquadrado numa escala numérica, que pode ser usado para fazer uma comparação objetiva entre ativos ou em relação a um valor de referência. Desta forma, esta métrica é uma importante contribuição para as decisões de manutenção que um operador tem que fazer.

Para calcular um Índice de Saúde de um ativo é necessário considerar os resultados de várias observações como o seu estado de operação, inspeções de campo, testes de diagnóstico e outras informações existentes sobre o ativo para obter um valor objetivo que represente a sua condição geral. Assim sendo, os principais passos a seguir para a formulação do Índice de Saúde são os seguintes (Dorison et al. 2007):

- Identificar os fatores que influenciam a condição do ativo;
- Estabelecer a correspondência entre os possíveis estados de cada fator para uma pontuação numa escala comum;
- Estimar a importância relativa de cada fator.

Um ativo pode ser composto por vários subsistemas, ou componentes, sendo necessário compreender bem as relações e dependências que estes têm entre si para que a sua condição possa ser avaliada e combinada num valor geral. Um ativo poder ter falhas apenas devido a um componente defeituoso que pode ser reparado ou substituído, não afetando o tempo de vida útil do ativo em si. No entanto, a degradação que um ativo sofre durante o seu envelhecimento tem que ser detetada de forma a diminuir o risco de este ter uma falha geral, potencialmente causando grandes danos. Um método de cálculo do Índice de Saúde bem desenvolvido deve quantificar esta degradação reconhecendo os diferentes tipos de falhas que o ativo pode ter e as suas causas.

De acordo com Hjartarson e Otal (2006), as propriedades essenciais de um Índice de Saúde são:

- Ser indicativo da capacidade do ativo para continuar em serviço e representativo da sua condição geral;
- Ser baseado em medidas objetivas e verificáveis do estado do ativo, em vez de observações subjetivas;
- Ser compreensível e facilmente interpretado.

O Índice de Saúde pode ser bastante útil na gestão de ativos. Se o seu processo de cálculo for bem formulado, a sua aplicação aos ativos existentes fornece uma indicação que deve ser considerada ao estabelecer prioridades, na alocação de recursos e a identificar possíveis oportunidades de investimento.

2.1.3.2 Probabilidade de Falha

De forma a avaliar a fiabilidade de um ativo é fundamental calcular a probabilidade de este ter uma falha. Este indicador representa a probabilidade de um ativo sofrer uma falha geral, tornando-o incapaz de cumprir a sua função, podendo não ser possível ou economicamente viável fazer a sua reparação. Além disso, a falha do ativo pode causar danos colaterais quer nos ativos que o rodeiam ou interagem com ele quer, a mais alto nível, nas entidades que dele dependem e, consequentemente, na reputação dos responsáveis pelo ativo.

Esta probabilidade pode ter uma grande correlação com o Índice de Saúde do ativo, uma vez que é esperado que a probabilidade de Falha de um ativo aumente com a sua degradação. Se o Índice de Saúde for formulado de forma correta e representar uma estimativa fiável da condição do ativo e dos seus componentes então este deve ter uma contribuição significativa no cálculo desta probabilidade. No entanto, há fatores externos que não são considerados no cálculo do Índice de Saúde que podem causar a falha do ativo. Por exemplo, as características do ambiente em que este está inserido como a poluição, risco de incêndio, presença de animais que possam interferir com o correto funcionamento do ativo, etc. Estes fatores têm que ser tidos em conta no cálculo da probabilidade de Falha.

Se houver um histórico detalhado das falhas que ocorreram nos ativos então este pode ser usado para calcular a distribuição probabilística dos tipos de falha e das suas causas, o que dá uma boa estimativa dos pesos a dar a cada fator tido em conta no cálculo da probabilidade de Falha. Desta forma, a importância individual das possíveis causas de falha pode ser refletida.

Tal como o Índice de Saúde, a probabilidade de Falha fornece um valor que é fácil de compreender, uma vez que é uma percentagem de 0 a 100, e que deve ser tido em conta no processo de decisão do operador. Além disso, pode ajudar a identificar ativos que tenham uma probabilidade de Falha elevada que não possam ser facilmente reparados ou substituídos, sendo importante aplicar medidas reativas mais sofisticadas de forma a aumentar a robustez do sistema em que o ativo está inserido, evitando maiores prejuízos.

2.1.3.3 Tempo de vida restante

O tempo de vida útil restante de um ativo é uma estimativa do tempo, tipicamente em anos, em que é expectável que o ativo continue a cumprir a sua função. Este indicador permite uma gestão mais a longo prazo, especialmente no planeamento de renovações e de medidas corretivas para abrandar o processo de envelhecimento de ativos.

É necessário estudar o ciclo de vida de cada tipo de ativo, tendo como base o histórico dos ativos desse tipo, os processos de degradação que estes sofrem, as medidas de manutenção preventiva aplicadas e outras informações como indicações do fabricante. Além do tipo de ativo, é preciso também fazer um estudo das variáveis que podem afetar o tempo de vida de um ativo em particular. Tal como no cálculo da probabilidade de Falha, os fatores externos a que um ativo está sujeito, como temperatura ambiente ou humidade, podem reduzir o seu tempo de vida em relação a outro ativo do mesmo tipo que esteja em condições menos severas. O histórico individual do ativo também é importante pois incidentes que tenham ocorrido e o tenham afetado temporariamente podem impatar o seu tempo de vida.

O cálculo do tempo de vida de um ativo tem que ter em conta a possibilidade de este se tornar obsoleto. Isto pode acontecer por várias razões:

- Alterações legais, leis que proíbam a utilização do ativo ou obriguem a modificações;
- Alterações das necessidades do proprietário;
- Motivos económicos, o ativo pode deixar de ser economicamente viável;
- Motivos tecnológicos, pode não ser possível reparar o ativo, por exemplo devido a deixarem de ser fabricadas as peças que um equipamento necessita.

Uma das grandes dificuldades deste tipo de análise é a rápida evolução tanto das características dos ativos como das técnicas de manutenção, o que reduz a significância do estudo do historial desse tipo de ativo. É essencial o conhecimento e experiência dos engenheiros para prever os efeitos das alterações no tempo de vida útil do ativo e fazer os ajustes necessários.

É fácil de perceber que estes três indicadores, o Índice de Saúde, a probabilidade de Falha e o tempo de vida restante, estão intimamente relacionados. Hjartarson e Otal(2006) sugerem uma relação entre os indicadores como se pode verificar na tabela 1.

Índice de Saúde	Condição	Probabilidade de Falha (PdF)	Posição na curva de vida	Requisitos
85-100	Muito Boa	Baixa	Primeira metade da vida expectável	Manutenção normal
70-85	Boa	Baixa mas aumentando ligeiramente	Segundo terço da vida expectável	Manutenção normal
50-70	Razoável	Aumentando rapidamente mas abaixo da PdF na idade média	Último terço da vida expectável	Aumento dos testes de diagnóstico. Atuar dependendo da criticidade
30-50	Má	Maior do que PdF na idade média e a aumentar	Primeiro terço depois da vida expectável	Planear processo de substituição/reparação considerando risco e consequência de falha
0-30	Muito Má	Muito alta, mais do dobro da PdF na idade média	Segundo terço depois da vida expectável	Avaliar risco e substituir/reparar caso necessário

Tabela 1 - Critérios para formulação do Índice de Saúde (Hjartarson and Otal 2006)

2.1.3.4 Risco

O risco de um ativo é a combinação da sua probabilidade de falha com as consequências de este falhar. É importante considerar o impacto que o ativo pode ter no sistema quando se faz o planeamento das intervenções de manutenção a efetuar ao longo do seu ciclo de vida. Esta análise permite controlar os riscos associados e justificar de forma mais fundamentada potenciais investimentos.

A consequência de uma falha pode ser classificada através de um valor numérico ou por níveis, por exemplo de “Negligenciável” a “Catastrófico” em que a falha provoque a perda de vida ou danos em pessoas. A consequência pode ser de vários tipos:

- Física: Danos colaterais, detioração acelerada, poluição, redução da robustez do sistema;
- Económica: Aumento dos custos a nível de manutenção e, possivelmente, devido a outras consequências como a contaminação do ambiente;
- Legais: Devido a multas e penalizações por danos provocados;
- Reputação: Repercussão do incidente na comunicação social e na comunidade;
- Social: Conflitos e incómodos gerados na população.

A relação entre a probabilidade de falha e as suas consequências pode ser resumida numa Matriz de Risco, uma ferramenta útil para comparar ativos e planear o seu processo de manutenção. Como exemplo, a figura 1 apresenta a Matriz de Risco utilizada na EDP Distribuição no Programa de Renovação e Reabilitação de Ativos.



Figura 1 - Matriz de Risco

2.2 Data Mining

O método manual de transformar dados em conhecimento assenta na sua análise e interpretação. Isto exige um analista com um conhecimento profundo sobre os dados e do seu domínio. Com o crescimento exponencial dos dados disponíveis também a complexidade de os analisar aumenta, tornando a tarefa demasiado difícil para um humano. De forma a responder às exigências de hoje, emergiram técnicas de processamento de enormes quantidades de dados para obter informações relevantes e conhecimento que possa ser compreendido por humanos.

Data Mining é o “processo não trivial de identificar padrões nos dados que sejam válidos, novos, potencialmente úteis e compreensíveis” (Fayyad et al. 1996). O conhecimento extraído tem que se manter verdadeiro quando aplicado em dados novos e resultar em benefícios para o utilizador. Devido ao seu sucesso em tirar partido dos dados disponíveis, o *Data Mining* tem uma grande aplicação em várias atividades como marketing, banca, deteção de fraude, telecomunicações, farmacêuticas ou na Internet.

2.2.1 Classificação e Regressão

As áreas de *Data Mining* com mais potencial para o problema em estudo são: Classificação e Regressão.

Um modelo de Classificação, ou classificador, tem o objetivo de prever a classe de um item. Esta classe é um rótulo, que pode ter valores discretos, por exemplo “Muito Bom” ou “Mau”. Um modelo de Regressão faz uma predição numérica, o seu resultado é um valor contínuo numérico (Han and Kamber 2006).

O processo de construir um modelo é similar para ambos os problemas. Os dados disponíveis têm que ser transformados em amostras, ou *samples*, que contêm os dados que o modelo recebe e o resultado esperado, aquele que é suposto ser previsto pelo modelo para essa amostra. As amostras disponíveis são divididas em conjunto de treino e em conjunto de teste. O conjunto de treino é usado para construir o modelo, este tenta detetar relações entre os dados de entrada e os resultados possíveis. Este passo é o que faz com que estes modelos sejam métodos de Aprendizagem Supervisionada.

De forma a validar a precisão do modelo obtido, é utilizado o conjunto de teste para comparar os resultados do modelo com os que realmente aconteceram. Para modelos de Classificação esta avaliação pode ser feita através de uma matriz de confusão que é uma ferramenta para comparar as classes previstas com as que se verificam na realidade, como demonstrado na tabela 2. A matriz pode ser utilizada para qualquer número de classes e identifica os verdadeiros positivos e negativos, ou seja, as predições corretas do classificador, e os falsos positivos e negativos, que são as classificações incorretas. Dependendo das características do problema, alguns dos valores podem ter maior importância, por exemplo num estudo de diagnóstico de uma doença um falso

negativo pode ser mais crítico que um falso positivo. A precisão do modelo é obtida dividindo a soma dos verdadeiros positivos e verdadeiros negativos pelo número total de predições.

		Classe prevista	
		Classe 1	Classe 2
Classe real	Classe 1	Verdadeiros positivos	Falsos negativos
	Classe 2	Falsos positivos	Verdadeiros negativos

Tabela 2 - Matriz de Confusão

No caso dos modelos de Regressão, como o valor previsto é um número exato, a precisão deve ser medida pela distância a que a predição ficou do valor observado. Métodos como o Erro Quadrático Médio medem a diferença média em todas as instâncias. Também pode ser utilizado o Erro da Raiz Quadrática Média, que tem o interesse de o erro estar na mesma magnitude dos valores a ser previstos.

Alguns dos algoritmos aplicados na construção destes modelos são: Árvores de Decisão, Redes Neurais, *Support Vector Machine*, *Random Forest*, entre outros.

O sucesso de um modelo de *Data Mining* pode ser influenciado por vários fatores. Os dados disponíveis podem não ser suficientes ou estar desequilibrados, quando há bastante mais casos de uma classe do que de outras, podendo influenciar o modelo a dar mais importância a essa classe. No entanto, é possível utilizar técnicas de balanceamento dos dados para anular estes desequilíbrios, como:

- *Up-Sampling*: Selecionar aleatoriamente amostras de cada classe para que tenham todas a frequência da classe com mais amostras;
- *Down-Sampling*: Selecionar aleatoriamente amostras de cada classe de forma a que tenham a mesma frequência da classe com menos amostras;

É fundamental analisar a capacidade preditiva de um modelo em amostras desconhecidas, que não tenham sido utilizadas no processo de treino, para verificar que este não está sobreajustado aos dados de treino, podem ser utilizados os seguintes métodos, entre outros, para o efeito:

- *K-fold Cross Validation*: As amostras são divididas aleatoriamente em K conjuntos, sendo um deles usado para teste e os outros para treino. É construído um modelo para cada um dos K conjuntos e são utilizadas as previsões feitas por cada um para avaliar a precisão geral.
- *Bootstrap*: São selecionadas aleatoriamente as amostras com repetição para treino, ou seja, uma amostra pode ser selecionada mais que uma vez. As amostras não selecionadas são o conjunto de teste.

2.3 Survival Analysis

Survival Analysis é um ramo da Estatística em que a variável objetivo é o tempo até que um determinado evento ocorra. Tem muitas aplicações na área da Saúde, por exemplo para estudar os efeitos de um medicamento no tempo de recuperação de um indivíduo e, também, em Engenharia para estudar a confiabilidade de um sistema ao longo do seu ciclo de vida.

Num problema de *Survival Analysis*, o tempo pode ser o intervalo de tempo desde o início do estudo até à ocorrência do evento no indivíduo ou a idade do indivíduo no momento em que ocorre o evento. O evento pode ser um acontecimento positivo ou negativo mas geralmente o tempo decorrido é referido como tempo de sobrevivência e o evento como falha. Se o tempo de falha de vários indivíduos é conhecido então esse conhecimento pode ser extrapolado para calcular várias métricas como o tempo médio de vida ou a percentagem esperada de indivíduos em que ocorra o evento até a um determinado tempo de vida. Por exemplo, quantos equipamentos com características semelhantes se espera que falhem antes dos dez anos de vida?

Um dos problemas comuns neste tipo de análise é que o evento ainda não tenha ocorrido numa parte da população à data de conclusão do estudo. Isto é especialmente prevalente no estudo de indivíduos com tempo de sobrevivência longo. Estes indivíduos são considerados como censurados à direita (Lawless, 1982). Um indivíduo também pode ser censurado à esquerda caso seja desconhecida a sua data de nascimento. Ao contrário de alguns métodos mais simples como a média ou a mediana, os métodos de *Survival Analysis* tentam tirar partido destes indivíduos censurados através das informações conhecidas até ao momento da censura.

Nesta secção, serão apresentados alguns conceitos essenciais para este tipo de estudo bem como alguns métodos usualmente utilizados.

2.3.1 Função de Sobrevivência

A função de Sobrevivência (*Survival function*) calcula a probabilidade de um indivíduo sobreviver até um determinado tempo t , como se pode verificar na seguinte fórmula:

$$S(t) = P(T > t) \quad (1)$$

Em que T é uma variável aleatória que pode ir de 0 a infinito e representa o tempo de sobrevivência do indivíduo. Uma função de Sobrevivência deve respeitar as seguintes regras (Kleinbaum and Klein 2012):

- Com o aumento do tempo t , $S(t)$ diminui;
- $S(0) = 1$, pois no início do estudo todos os indivíduos estão vivos, ou seja, o evento ainda não ocorreu;
- $S(\infty) = 0$, se o tempo do estudo não tiver limite então é esperado que todos os indivíduos morram.

2.3.2 Função de Risco

A função de Risco (*Hazard function*) dá a taxa de falha condicional de um indivíduo num determinado tempo t , ou seja, o potencial instantâneo por unidade de tempo de que o evento ocorra dado que o indivíduo sobreviveu até ao tempo t (Kleinbaum and Klein 2012), como demonstrado na seguinte fórmula:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \quad (2)$$

Se a função de Risco for conhecida a função de Sobrevivência pode ser deduzida a partir desta através da fórmula 3, e vice-versa, usando a fórmula 4:

$$S(t) = \exp \left\{ - \int_0^t h(u) du \right\} \quad (3)$$

$$h(t) = - \frac{1}{S(t)} \frac{dS(t)}{dt} \quad (4)$$

2.3.3 Kaplan-Meier

Os métodos de *Survival Analysis* não paramétricos permitem estimar as funções de Sobrevivência e de Risco de uma população sem que seja necessário assumir uma distribuição estatística dos dados de sobrevivência. Além disso, suportam populações que contenham indivíduos censurados.

O estimador Kaplan-Meier (Kaplan and Meier 1958) é o método Não Paramétrico mais comum para este tipo de problema. A função de Sobrevivência é calculada segundo a seguinte fórmula:

$$\hat{S}(t) = \prod_{t_i < t} \frac{n_i - d_i}{n_i} \quad (5)$$

Os tempos distintos em que ocorrem os eventos conhecidos são ordenados de $t(1)$ ao tempo t . Sendo d_i o número de indivíduos em que o evento ocorreu no tempo t_i e n_i o número de indivíduos vivos antes de t_i .

Este método tem a vantagem de apenas precisar dos tempos de falha de cada indivíduo para estimar uma função de Sobrevivência. No entanto, isto também significa que o cálculo só tem em conta a idade dos indivíduos ignorando as suas outras propriedades e condição. Ainda assim é um método bastante útil para estimar uma curva de sobrevivência para servir de referência.

2.3.4 Cox Proportional Hazards

O modelo *Cox Proportional Hazards* é um método Semi-Paramétrico e expressa a função de Risco de uma população através da seguinte fórmula:

$$h(t, X) = h_0(t) e^{\sum_{i=1}^p \beta_i X_i} \quad (6)$$

X é um conjunto de variáveis explanatórias a ser modeladas para prever o risco de ocorrer o evento num indivíduo. A fórmula é dividida em duas partes, a primeira parte, $h_0(t)$, é a função de Risco de referência e tem apenas em conta o tempo t , não fazendo uso das variáveis de X . A segunda parte não envolve o tempo t mas utiliza os valores correspondentes ao indivíduo em causa, através das variáveis de X . Estas variáveis são independentes em relação ao tempo como, por exemplo, o sexo de um indivíduo. É por esta razão que o modelo *Cox Proportional Hazards* é um modelo Semi-Paramétrico, uma vez que caso não existam variáveis explanatórias o resultado do modelo será igual à função de risco de referência que apenas tem em conta o tempo.

Este método é popular devido à sua robustez uma vez que é capaz de ter resultados aproximados a um modelo paramétrico que se adeque aos dados. Caso seja conhecida a distribuição estatística dos dados em análise um modelo paramétrico é a melhor opção, mas em caso de dúvida o modelo de Cox é uma escolha segura (Kleinbaum and Klein 2012).

2.3.5 Métodos Paramétricos

Alguns métodos baseados em distribuições probabilísticas podem ser aplicados em problemas de *Survival Analysis*. Assume-se que a variável objetivo, que neste caso é o tempo de sobrevivência, segue uma distribuição cuja função de densidade pode ser expressa por parâmetros desconhecidos. O objetivo é otimizar os parâmetros da distribuição de forma a ajustá-los aos dados, utilizando técnicas como o método *Least Squares* ou *Maximum-Likelihood estimation*. Conhecendo a função de densidade, $f(t)$, é possível obter as funções de Sobrevivência, $S(t)$, e de Risco, $f(t)$, através das equações 7 e 8.

$$S(t) = P(T > t) = \int_0^t f(u)du \quad (7)$$

$$f(t) = h(t)S(t) \quad (8)$$

Desta forma, as leis e princípios que afetam a população em estudo, e que não são diretamente observáveis, podem ser formuladas em termos de hipótese (Cota-Felix et al. 2009). O processo deste tipo de método é o seguinte (Maciejewski et al. 2011):

- Selecionar uma distribuição;
- Otimizar os parâmetros da distribuição;
- Testar o grau de ajuste aos dados conhecidos;
- Se o ajuste for satisfatório, derivar as funções de Sobrevivência e de Risco.

Os modelos paramétricos têm ainda a vantagem de especificar completamente as curvas de Sobrevivência e de Risco, sendo estas mais consistentes com as curvas teóricas. Já nos métodos de Kaplan-Meier e de Cox as curvas são em forma de degraus e, especialmente quando na população em estudo há poucos indivíduos não censurados, pode haver grandes intervalos de tempo entre duas estimativas. Esta diferença pode ser observada comparando os gráficos das figuras 2 e 3:

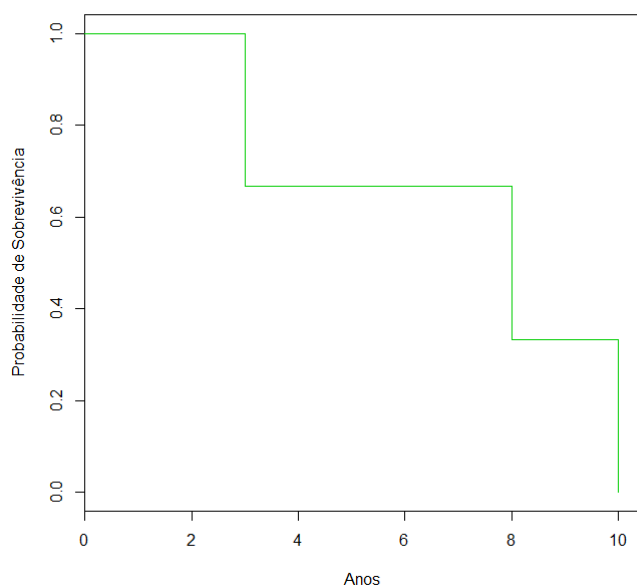


Figura 2 - Curva de Sobrevivência Kaplan-Meier

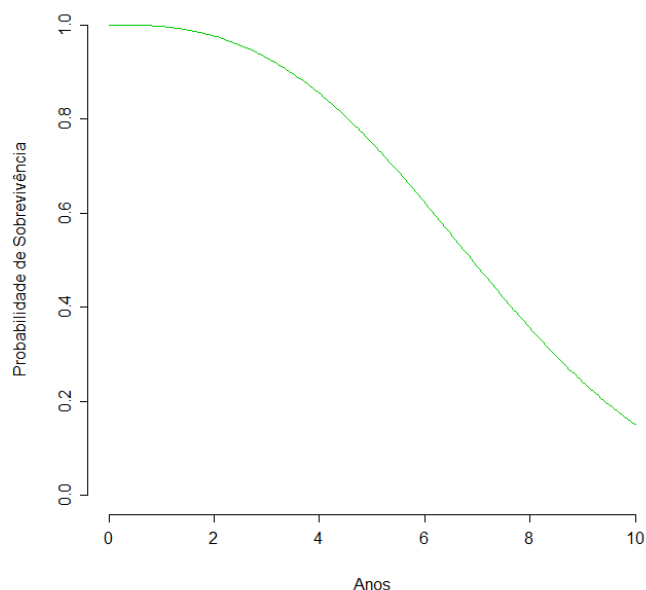


Figura 3 - Curva de Sobrevivência Weibull

O gráfico da Figura 2 representa a curva de Sobrevivência gerada pelo método Kaplan-Meier para uma população com poucos indivíduos e o da Figura 3 representa a curva gerada através de uma distribuição de Weibull para a mesma população.

Testar o grau de ajuste pode ser problemático se forem conhecidos poucos indivíduos não censurados pois os critérios mais objetivos usam esses indivíduos.

2.3.5.1 Distribuição de Weibull

A distribuição de Weibull é umas das distribuições que pode ser utilizada para estimar a curva de Sobrevivência de uma população. As funções de Sobrevivência e de Risco são calculadas através das seguintes fórmulas:

$$S(t) = \exp(-\lambda t^\rho) \quad (9)$$

$$h(t) = \lambda \rho t^{\rho-1} \quad (10)$$

A função de densidade é igual ao produto das duas funções, como expresso na equação 8. O parâmetro *shape*, representado por ρ , determina a forma da função de Risco, permitindo à distribuição de Weibull ter alguma flexibilidade:

- Se $\rho > 1$, o risco aumenta com o passar do tempo;
- Se $\rho = 1$, o risco mantém-se constante;
- Se $\rho < 1$, o risco diminui com o passar do tempo.

O parâmetro λ representa o risco constante e é calculado tendo em conta as variáveis independentes conhecidas nos indivíduos em estudo. Por exemplo, num estudo em que haja uma variável independente X_1 , λ é calculado através da fórmula 11, sendo feita uma estimativa dos parâmetros β_0 e β_1 para ajustar a distribuição aos dados.

$$\lambda = \exp(\beta_0 + \beta_1 X_1) \quad (11)$$

2.3.5.2 Distribuição Exponencial

A distribuição Exponencial é um caso especial da distribuição de Weibull quando $\rho = 1$. Isto significa que nesta distribuição o risco mantém-se constante durante o tempo de vida dos indivíduos. Assim sendo, as fórmulas das funções de Sobrevivência e de Risco são uma simplificação das equações 9 e 10:

$$S(t) = \exp(-\lambda t) \quad (12)$$

$$h(t) = \lambda \quad (13)$$

$$f(t) = \lambda \exp(-\lambda t) \quad (14)$$

2.3.6 Árvore de Decisão

Uma Árvore de Decisão é um modelo com uma estrutura em forma de árvore que é construída a partir dos atributos dos indivíduos em estudo para prever a variável objetivo. Os nós representam um teste a uma variável, por exemplo se a idade de um indivíduo é maior que 20 anos, e os ramos são os possíveis resultados desse teste. Os nós terminais, ou folhas, representam uma classificação, indicando que os indivíduos correspondentes a essa folha pertencem a uma classe da variável objetivo. Este processo é feito de forma recursiva e o modelo tenta selecionar os melhores testes (*splits*) que são os que representam um ganho de informação maior, ou seja, são capazes de dividir os indivíduos em grupos mais homogêneos.

Uma Árvore de Decisão pode crescer demasiado, podendo até gerar um nó final para cada indivíduo. Isto pode fazer com que o modelo fique demasiado adaptado aos dados de treino e tenha baixa precisão em dados que não conheça. Para contrariar isto, é necessário podar a árvore, através de técnicas de *Pruning*. Pode ser determinado um número mínimo de indivíduos que um nó final tem que ter ou cortar nós finais minimizando a perda de precisão do modelo.

Uma das vantagens deste modelo é ser fácil de interpretar por um humano, o que facilita a sua validação e possibilita que se extraiam regras sobre os indivíduos. Além disso, é capaz de suportar vários tipos de variáveis, como textuais, numéricas, discretas ou contínuas. No entanto, apenas consegue fazer previsões em que as variáveis independentes estejam dentro dos limites no conjunto de treino.

2.3.7 Random Survival Forest

Random Survival Forest (Ishwaran et al. 2008) é uma adaptação do método *Random Forest* (Breiman 2001) para análise de dados censurados à direita, podendo ser aplicado em problemas de *Survival Analysis*. É um método *ensemble*, que se baseia em construir múltiplas Árvore de Decisão de forma aleatória. Esta aleatoriedade é introduzida de duas formas, cada Árvore utiliza uma amostra aleatória diferente dos dados disponíveis e as variáveis consideradas em cada *split* são escolhidas aleatoriamente, garantindo que haja diferenças entre as várias Árvore de Decisão. Como o problema de *Survival Analysis* se enquadra numa Regressão, a predição do modelo é igual à média das predições feitas por todas as árvores. Este método demonstra grande capacidade de previsão, mantendo um erro de generalização baixo (Ishwaran et al. 2008).

Tanto os métodos paramétricos apresentados anteriormente como o modelo *Cox Proportional Hazards* requerem que as variáveis independentes a considerar sejam escolhidas à partida, exigindo uma análise extensiva em casos onde hajam múltiplas variáveis para comparar todas as combinações possíveis ou requer um conhecimento extenso sobre o problema para facilitar essa tarefa, introduzindo subjetividade. A aplicação de *Random Forests* permite que esse processo seja automático.

Os passos para criar o modelo são os seguintes:

- São extraídas B amostras por *Bootstrap*;
- É construída uma Árvore de Decisão para cada amostra. Em cada nó, são selecionadas p variáveis para a divisão dos indivíduos, que é feita maximizando a diferença de tempo de sobrevivência dos indivíduos em cada nó filho;
- A árvore é construída enquanto for cumprido o requisito de os nós terminais terem um número mínimo de mortes, definido previamente;
- Os dados que não foram usados para construir as árvores são utilizados para verificar o erro de previsão do modelo.

Para que o modelo suporte indivíduos censurados é utilizado um estimador Não Paramétrico, o método *Nelson-Aalen*, que estima a função de Risco Cumulativo através da seguinte equação:

Em que é d_i o número de eventos em t_i e n_i o número de indivíduos vivos imediatamente antes desse momento. Esta função dá o número cumulativo de mortes até um determinado

$$\tilde{H}(t) = \sum_{t_i \leq t} \frac{d_i}{n_i} \quad (15)$$

momento. Esta estimação é feita em cada nó da Árvore, sendo que dentro de um nó terminal todos os indivíduos têm a mesma função. No final do processo, a função de Risco Cumulativo do modelo é igual à média da função de todas as árvores

Capítulo 3

Gestão de Ativos numa subestação elétrica

A EDP Distribuição é responsável por uma extensa Rede de Distribuição e tem como função fazer chegar a energia elétrica gerada nos centros de produção aos seus clientes. É necessário garantir o bom funcionamento dos ativos envolvidos para minimizar o número de interrupções e atingir uma qualidade de serviço que satisfaça os consumidores.

Uma subestação elétrica é uma infraestrutura constituída por vários equipamentos cuja principal função é transformar a tensão da corrente elétrica na rede. O tipo de subestação em estudo nesta dissertação é a subestação AT/MT (Alta e Média Tensão). Estas subestações transformam a eletricidade produzida nas centrais elétricas, transportada em Alta Tensão para minimizar as perdas de eletricidade, de Alta para Média Tensão. Posteriormente, a corrente é transformada de Média para Baixa Tensão nos transformadores locais perto das zonas de habitação, podendo assim ser utilizada pelos consumidores finais. A rede de distribuição nacional conta com cerca de 450 Subestações. Uma subestação é constituída pelo edifício de comando onde se encontra o equipamento de Média Tensão, os sistemas de alimentação e SPCC, com regulação de temperatura e deteção de incêndio e de intrusos. No exterior há um parque de



Figura 2 - Edifício de comando



Figura 3 - Parque de aparelhagem

aparelhagem onde são instalados os equipamentos de Alta Tensão e equipamentos complementares.

Com o aumento do consumo energético é essencial modernizar e reforçar a rede de distribuição. Além disso, com os avanços na indústria como a melhoria das técnicas de manutenção e evolução tecnológica dos equipamentos, a idade média dos ativos tem vindo a aumentar tornando-os mais suscetíveis a falhas.

Em 2004, um estudo feito nos Estados Unidos da América estimou que, anualmente, os prejuízos dos consumidores causados por cortes de energia cheguem aos 80 mil milhões (LaCommare and Eto 2004). Para minimizar as interrupções devido a falhas, respeitando o orçamento disponível e os objetivos da empresa, é fundamental otimizar o processo de gestão dos ativos envolvidos, identificando os fatores de risco e atuando antes que ocorram as falhas. Além das consequências para o consumidor final, a falha de um ativo pode causar danos nos restantes equipamentos e no ambiente que o rodeia gerando elevados prejuízos para a empresa. Os custos das intervenções de manutenção corretiva, para reparar um ativo e voltar a colocá-lo em serviço, são geralmente mais elevados do que intervenções preventivas que evitem que este falhe.

A transição para um modelo de manutenção preditiva exige um conhecimento profundo dos processos de degradação dos ativos e capacidade de caracterizar a sua condição de forma fiável para planear e programar as intervenções de manutenção de acordo com o estado dos ativos. Para responder às necessidades atuais, foi criado o projeto “subestação 5M” com os seguintes objetivos:

- Subestações Mais Simples;
- Subestações Mais Pequenas;
- Subestações Mais Económicas;
- Subestações Mais Fáceis de Manter;
- Subestações Rápidas de Construir.

No âmbito deste projeto, com foco nos objetivos de tornar as subestações mais económicas e fáceis de manter, foram desenvolvidos algoritmos para caracterizar a condição dos principais ativos de uma subestação. Estes algoritmos estimam os seguintes indicadores:

- Índice de Saúde;
- Probabilidade de Falha;
- Tempo de vida restante.

Através destes indicadores pretende-se reunir as condições necessárias para analisar os ativos de uma forma objetiva, permitindo a priorização das intervenções de manutenção e a otimização da gestão das subestações tendo sempre em vista a qualidade de serviço e consequente satisfação do cliente.

Neste capítulo são apresentados os principais ativos de uma subestação e os algoritmos desenvolvidos na EDP para estimar os indicadores referidos. Na última secção é relatada uma análise de ensaios em disjuntores que permitiu recolher informações sobre estes e melhorar o seu algoritmo de avaliação.

3.1 Principais Ativos de uma subestação

Tendo em conta a sua função e importância, foram identificados os seguintes equipamentos como os principais ativos de uma subestação elétrica:

- Transformadores;
- disjuntores;
- Sistemas de Proteção Comando e Controlo;
- Sistemas de Alimentação.

Nesta secção serão apresentados estes ativos, caracterizando-os e explicitando as suas funções.

3.1.1 Transformador de Potência

Os Transformadores de Potência são o equipamento mais valioso numa subestação elétrica, significando quase 60% do investimento total (Jahromi et al. 2009). Além do seu valor, a sua função é imprescindível e são necessárias medidas de segurança devido ao elevado perigo em caso de falha. Como tal, é vital aferir a sua condição para evitar que falhe e se danifique de forma irreparável. Um planeamento correto das intervenções de manutenção permite aumentar o tempo de vida útil de um transformador e rentabilizar ao máximo este importante ativo.

A função de um Transformador de Potência numa subestação AT/MT é transformar a corrente de Alta Tensão (60 kV) para Média Tensão (30, 15 ou 10kV) garantindo a potência do circuito. Os enrolamentos do transformador são imersos em óleo com papel que serve de isolamento e protege-os de pressões térmicas, elétricas e mecânicas.

De acordo com o relatório da CIGRE sobre a caracterização do estado de transformadores de potência (Hhlein et al. 2003), o tempo de vida e a degradação de um transformador são afetados principalmente pelas seguintes propriedades:

- Dielétricas, o isolamento do transformador, por exemplo nos enrolamentos;
- Magnéticas, a condição do núcleo e da estrutura que o envolve;
- Comutador, um mecanismo que permite a regulação da voltagem do transformador;
- Integridade mecânica, a degradação que os vários componentes do transformador sofrem em condições de operação, como a cuba, o sistema de refrigeração, etc.

Na figura 2 é apresentado um transformador, este pode pesar até cerca de 70 toneladas e atingir cerca de 5 metros de altura. O óleo isolante compõe usualmente 25% do peso total.



Figura 2 - Transformador de Potência

3.1.2 Disjuntor

O disjuntor é um dispositivo eletromecânico que tem como função permitir manobras em carga e efetuar o corte de energia quando ocorrem curto-circuitos ou sobrecargas. Quando ocorre uma falha é necessário interromper a corrente para evitar que esta se propague para os restantes equipamentos e os danifique. Desta forma o circuito elétrico é interrompido antes que as pressões térmicas e mecânicas se tornem prejudiciais.

O tipo de disjuntor em estudo é trifásico, possuindo três polos como se pode verificar na figura 3. Em cada polo estão dois pontos de contacto que são separados quando é necessário efetuar um corte. Os polos têm que suportar elevadas pressões quando o corte ocorre, pois é gerado um arco de eletricidade que tem de ser extinto. Este processo deverá ter uma duração muito curta e há varias tecnologias para o conseguir:

- Óleo, imergindo os pontos de contacto para suprimir o arco elétrico;
- Ar comprimido, extinguindo o arco através de jatos de ar;
- Vácuo, não havendo matéria para o arco ionizar sem ser o material isolante;

- SF6, utilizando o gás hexafluoreto de enxofre (SF6) como meio isolante e de extinção do arco.



Figura 3 - Disjuntor

Numa subestação os disjuntores de Média Tensão encontram-se dentro do edifício de comando numa Cella de Média Tensão onde estão blindados e isolados ao ar como se pode verificar na figura 4.



Figura 4 - Cella de Média Tensão

3.1.3 Sistemas de Proteção, Comando e Controlo

Os Sistemas de Proteção, Comando e Controlo, ou SPCC, são um conjunto de dispositivos que têm o objetivo de proteger, monitorizar e permitir o comando e controlo dos equipamentos de uma subestação localmente e à distância.

São utilizados sensores e atuadores para permitir a deteção em tempo real de defeitos no sistema elétrico e prevenir a ocorrência de falhas.

São implementadas várias opções para comunicar com a subestação:

- Fibra ótica, com elevada fiabilidade e largura de banda;
- Comunicação *Power Line* (PLC), que utiliza a rede de energia como meio de transmissão;
- Via radiofrequência (FR).

Através das vias de comunicação é possível os SPCC transmitirem as medidas recolhidas e receberem comandos para efetuar manobras nos equipamentos da subestação de forma remota, sem exigir a presença de técnicos no local.

Um dos equipamentos fundamentais dos SPCC é o IED (Intelligent Equipment Device), apresentado na figura seguinte.



Figura 5 - Intelligent Equipment Device

Cada painel da subestação está ligado a um destes dispositivos que são responsáveis pela sua proteção e supervisão. Por exemplo, se detetar uma variação perigosa na corrente elétrica envia um comando ao disjuntor para que este interrompa o circuito. Estas unidades têm funções de autodiagnóstico e permitem maior fiabilidade dos equipamentos. Além disso, fazem o registo das ocorrências no sistema, como comandos falhados ou medidas não adquiridas. Os vários IEDs

de uma subestação estão interligados por rede de fibra ótica permitindo funções de automatismo distribuídas.

Uma subestação tem ainda uma Unidade Central de Processamento que reúne as informações de cada painel como a listagem de medidas e de automatismos. Esta unidade fornece uma interface para controlo local permitindo a um operador controlar o estado dos equipamentos.



Figura 6 - Unidade Central de Processamento

3.1.4 Sistemas de Alimentação

Os Sistemas de Alimentação são constituídos por Alimentadores e Baterias e têm como função garantir a alimentação dos sistemas auxiliares de uma subestação em caso de falha da alimentação em corrente alternada.

Os equipamentos de comando e proteção e outros serviços como a detecção de incêndios e intrusos são alimentados por tensão contínua através de baterias. Numa situação de funcionamento normal, uma subestação é alimentada em tensão alternada que é convertida em corrente contínua para carregar as baterias. Caso ocorra um corte da corrente a energia armazenada nas baterias assegura o comando funcional da subestação, permitindo o funcionamento das proteções e o comando dos equipamentos. Normalmente, as baterias têm uma autonomia de 10 horas.

Estes sistemas de alimentação são monitorizados e têm funções de autodiagnóstico e de manutenção da vida útil das baterias.



Figura 7 - Baterias de Tensão Contínua

3.2 Algoritmos desenvolvidos na EDP

Nesta secção serão apresentados os algoritmos desenvolvidos anteriormente pela empresa e a metodologia adotada para calcular os indicadores Índice de Saúde, Probabilidade de Falha e tempo de vida restante para cada tipo de ativo. Para complementar a explicação é apresentada como exemplo a adaptação dos algoritmos para o caso do disjuntor. A aplicabilidade destes indicadores no processo de gestão de uma subestação elétrica é também discutido.

3.2.1 Índice de Saúde

É pretendido que o Índice de Saúde seja indicativo da condição técnica do ativo tendo como base as suas características. Deve ser calculado da forma mais objetiva possível para que ativos iguais sejam comparados nas mesmas condições.

Foi adotada uma escala de 0 a 100 para o Índice de Saúde, sendo que 100 representa um ativo em condições ótimas e 0 um ativo em estado crítico.

Cada um dos ativos identificados foi decomposto em componentes que representam blocos funcionais e são responsáveis por assegurar uma ou mais funções. Para cada função foram identificadas as possíveis falhas que impeçam o ativo de a cumprir. Por fim, são associados um grupo de parâmetros a cada falha. A seguinte figura ilustra esta estrutura:

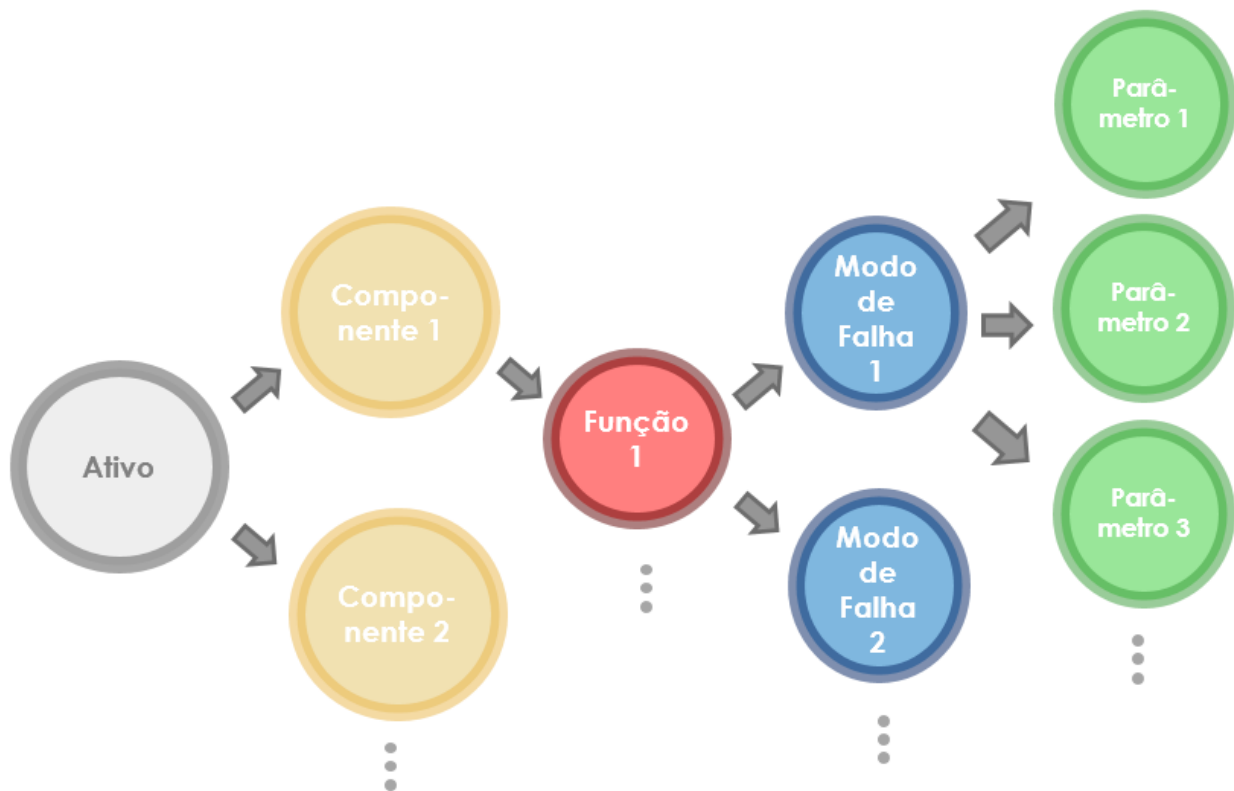


Figura 8 - Estrutura do algoritmo

Um parâmetro representa uma propriedade mensurável do ativo, que pode ser obtido através de diagnósticos e ensaios, inspeções visuais ou termográficas, monitorização *online* ou sinais de alarme que podem ser despoletados. O método de aquisição de um parâmetro é devidamente identificado.

O valor de um parâmetro é convertido para um Índice de Saúde através de uma das seguintes formas:

- Por escalões, em que os valores possíveis são divididos em intervalos que por sua vez correspondem a um Índice de Saúde;
- Por uma função, que recebe o valor do parâmetro e retorna o Índice correspondente. Esta alternativa permite uma valorização mais precisa do parâmetro.

Dependendo da natureza do ativo, este pode ter uma de várias tecnologias possíveis o que implica que alguns parâmetros tenham valores de referência diferentes e que nem todos os parâmetros se apliquem a todas as tecnologias.

Há ainda parâmetros globais que não pertencem a um modo de falha específico mas influenciam o Índice de Saúde de todos os componentes. Estes parâmetros são características mais gerais do ativo como a sua idade ou atrasos que tenham ocorrido na sua manutenção.

Para calcular a contribuição de um modo de falha para o Índice de Saúde é feito um majorante dos parâmetros que lhe estão associados, sendo escolhido o pior caso. Isto impede que um parâmetro que indique que o ativo está num estado degradado seja anulado pelos restantes parâmetros que não identificam essa degradação.

A cada modo de falha são associados um peso global e um peso específico. Estes pesos são percentagens que são utilizadas para valorizar a contribuição desse modo de falha para o Índice de Saúde do componente de que este faz parte, no caso do peso específico, e para o Índice de Saúde Global do ativo, através do peso global. A soma dos pesos dos modos de falha e parâmetros envolvidos no cálculo do Índice é igual a 100%.

Uma das situações a considerar é não ser possível adquirir o valor de um determinado parâmetro. Isto pode acontecer devido a características específicas do ativo ou por falta de recursos. Este fator é introduzido no algoritmo através de um indicador de Confiança que reflete o grau de certeza no valor Índice de Saúde face aos parâmetros conhecidos. Quando um valor não é conhecido o seu peso não é considerado o que faz com que o peso total seja inferior a 100%. O Índice de Saúde é dividido pelo peso total para que seja normalizado em relação aos outros ativos, devendo o gestor responsável julgar se o nível de confiança é suficiente para representar o estado do ativo.

Podem ser atribuídos avisos e alarmes aos parâmetros para que estes sejam despoletados quando são atingidos valores preocupantes ou críticos. Esta funcionalidade em conjunto com a atribuição de um Índice de Saúde específico a cada componente permite identificar os fatores com mais impacto no Índice de Saúde global do ativo. Isto simplifica a perceção das necessidades

específicas do ativo e auxilia o gestor no processo de decisão das medidas de manutenção a aplicar.

A organização modular do algoritmo facilita a incorporação de novos parâmetros, que surjam com a evolução das tecnologias de monitorização e do *know-how* dos operadores. Além disso, é expectável que, com o tempo de utilização deste tipo de manutenção, a estimação da condição seja refinada, sendo relativamente simples ajustar o cálculo do Índice de Saúde através da valorização dos parâmetros e do peso atribuído a cada modo de falha.

Na tabela 3 é apresentada a decomposição de um disjuntor em componentes, as suas funções e correspondentes modos de falha.

Componente	Função	Modo Falha
Câmara de corte	Garantir Estabelecimento Cir. Elétrico Garantir Isolamento Circ. Elétrico	Perda de isolamento interno
		Perda de isolamento externo
Comando	Garantir Estabelecimento Circ. Elétrico.	Falha operação ligar
	Garantir Isolamento Circ. Elétrico.	Falha operação desligar
	Garantir Estabelecimento Circ. Elétrico. Garantir Isolamento Circ. Elétrico.	Falha operação ligar- Falha operação desligar.

Tabela 3 - Modos de falha de um disjuntor

Como se pode verificar, o disjuntor é dividido em “Câmara de corte” e “Comando”, tendo sido identificados cinco modos de falha. Na tabela 4 são identificados os parâmetros que impactam o Índice de Saúde da “Câmara de corte” e na tabela 5 estão os parâmetros relativos ao Índice de Saúde do “Comando”. Para cada parâmetro estão identificados os níveis previstos e o seu método de aquisição. Também é apresentado para cada modo de falha o seu peso específico e global.

Os parâmetros “Idade” e “Atraso após a data da manutenção prevista” que são parâmetros globais, afetando ambos os componentes. De destacar o parâmetro “Resistência de Contacto” que é um exemplo de um parâmetro cujos níveis variam de acordo com a tecnologia do ativo.

Componente	Função	Modo Falha	Parâmetro	Tecnologia Disjuntor	Graus de Valorização					Majorante	Ponderações IS		Aquisição		
					Nível 0	Nível 1	Nível 2	Nível 3	Crítico		Global	Bloco Funcional			
			Idade	Todas	Função				-	NA	10.0%	10.0%	Gestor Activos		
Câmara de corte	Garantir Estabelecimento Círc. Eléctrico Garantir Isolamento Círc. Eléctrico	Perda de isolamento interno	Corrente cortada	Todas	$\Sigma < 0,9$ limite máx.			limite máx> $\Sigma \geq 0,9$ limite máx.	$\Sigma \geq$ limite máx.	X	22.5%	40.0%	SPCC		
			Resistência Contacto	SF6	$RC \leq 50 \mu\Omega$	$50 \mu\Omega \leq RC \leq 200 \mu\Omega$	$200 \mu\Omega \leq RC \leq 400 \mu\Omega$	$400 \mu\Omega \leq RC \leq 600 \mu\Omega$	$RC > 600 \mu\Omega$				Ensaio local		
				Óleo	$RC \leq 100 \mu\Omega$	$100 \mu\Omega \leq RC \leq 400 \mu\Omega$	$400 \mu\Omega \leq RC \leq 700 \mu\Omega$	$700 \mu\Omega \leq RC \leq 1000 \mu\Omega$	$RC > 1000 \mu\Omega$						
				Vácuo	$RC \leq 40 \mu\Omega$	$40 \mu\Omega \leq RC \leq 60 \mu\Omega$	$60 \mu\Omega \leq RC \leq 100 \mu\Omega$	$100 \mu\Omega \leq RC \leq 200 \mu\Omega$	$RC > 200 \mu\Omega$						
			Resistência Isolamento	Todas	$RI \geq 1 G \Omega$		$0,5 G \Omega < RI < 1 G \Omega$	$RI < 0,5 G \Omega$	Ensaio local						
			Fugas do meio de corte (PVO e SF6)	PVO e SF6	Sem fuga		Alarme	Abaixo do nível mínimo	SF6 - TC PVO - Inspeção Visual						
			Termografia (PQ interno)	Todas	PQ = 0	$PQ C \geq 1$	$PQ B \geq 1$	$PQ A \geq 1$	$PQ \geq 100 ^\circ C$				Termografia		
		Perda de isolamento externo	Resistência Isolamento	Todas	$RI \geq 1 G \Omega$		$0,5 G \Omega < RI < 1 G \Omega$	$RI < 0,5 G \Omega$	X	22.5%	40.0%	Ensaio local			
			Estado dos isoladores (fissuras e contornamentos)	Todas	Sem danos	Sujidade	Fugas de óleo / Isoladores lascados	Fissuras / contornamentos				Isolador partido	Inspeção Visual		
			Termografia (PQ externo)	Todas	PQ = 0	$PQ C \geq 1$	$PQ B \geq 1$	$PQ A \geq 1$				$PQ \geq 100 ^\circ C$	Termografia		
					Atraso após a data da manutenção prevista	Todas	Atraso até 20%	Atraso até 50%	Atraso até 80%	Atraso > 80%	-	NA	10.0%	10.0%	SAP-PM

Tabela 4 - Parâmetros impactantes no Índice de Saúde da Câmara de corte

Componente	Função	Modo Falha	Parâmetro	Tecnologia Disjuntor	Graus de Valorização					Majorante	Ponderações IS		Aquisição	
					Nível 0	Nível 1	Nível 2	Nível 3	Crítico		Global	Bloco Funcional		
Comando	Garantir Estabelecimento Círc. Eléctrico.	Falha operação ligar	Idade	Todas	Função					-	NA	10.0%	10.0%	Gestor Activos
			Tempo Fecho	SF6	65ms ≥ TF	65ms < TF ≤ 90ms	90ms < TF ≤ 120ms	120ms < TF ≤ 180ms	TF > 180ms	NA	5.0%	10.0%	Ensaio local / SPCC	
				Óleo	65ms ≥ TF	65ms < TF ≤ 90ms	90ms < TF ≤ 120ms	120ms < TF ≤ 180ms	TF > 180ms					
	Garantir Isolamento Círc. Eléctrico.	Falha operação desligar	Tempo Abertura	SF6	50ms ≥ TA	50ms < TA ≤ 80ms	80ms < TA ≤ 120ms	120 ms < TA ≤ 180 ms	TA > 200ms	NA	15.0%	35.0%	Ensaio local / SPCC	
				Óleo	50ms ≥ TA	50ms < TA ≤ 80ms	80ms < TA ≤ 120ms	120 ms < TA ≤ 180 ms	TA > 200ms					
				Vácuo	50ms ≥ TA	50ms < TA ≤ 80ms	80ms < TA ≤ 120ms	120 ms < TA ≤ 180 ms	TA > 200ms					
	Garantir Estabelecimento Círc. Eléctrico. Garantir Isolamento Círc. Eléctrico.	Falha operação ligar- Falha operação desligar.	Nº Manobras	SF6	NM ≤ 2000	2500 < NM ≤ 6000	6000 < NM ≤ 10000	NM > 10000	-	X	15.0%	35.0%	Inspeção Visual	
				Óleo	NM ≤ 2500	2500 < NM ≤ 6000	6000 < NM ≤ 10000	NM > 10000	-					
				Vácuo	NM ≤ 2500	2500 < NM ≤ 6000	6000 < NM ≤ 10000	NM > 10000	-					
			Estado das tubagens de acoplamento aos polos ou existência de fugas	Todas	Sem Degradação	Degradação ligeira	Pontos Corrosão	Degradação geral ou com fugas	-				Inspeção Visual	
Estado dos acoplamentos mecânicos comando aos polos			Todas	Sem Degradação	Degradação ligeira	Pontos Corrosão	Degradação geral	-						
Atraso após a data da manutenção prevista				Atraso até 20%	Atraso até 50%	Atraso até 80%	Atraso > 80%	-	NA				10.0%	10.0%

Tabela 5 - Parâmetros impactantes no Índice de Saúde do Comando

3.2.2 Probabilidade de Falha

A probabilidade de falha de um ativo indica a probabilidade de o ativo parar de desempenhar a sua função. Apesar de estar relacionado com o Índice de Saúde Visto que representa uma probabilidade, este indicador é uma percentagem de 0 a 100 por cento.

Como foi visto anteriormente, os ativos são divididos em componentes, sendo identificados os seus possíveis modos de falha. Para cada um destes modos é estimada uma probabilidade de falha e a probabilidade geral de o ativo falhar incorpora essas percentagens.

No algoritmo desenvolvido são tidos em conta quatro componentes para o cálculo da probabilidade de um modo de falha:

- As características técnicas do ativo, normalmente a sua idade;
- O Índice de Saúde associado ao modo de falha;
- Os Fatores externos, relativos à localização do ativo como o risco sísmico ou poluição, que impactam o modo de falha;
- Condições de operação a que o ativo esteve sujeito.

A cada um destes componentes é atribuído um peso percentual que representa a sua contribuição para a probabilidade de falha.

Estimadas as probabilidades de falha de todos os modos de falha, é necessário calcular a probabilidade de falha geral do ativo. Para tal é utilizado o princípio da inclusão-exclusão. Considerando que as diversas falhas não são eventos mutuamente exclusivos, por exemplo a perda de isolamento externo de um disjuntor não impede a perda de isolamento interno, a probabilidade de acontecer uma falha é igual à união de todas as falhas possíveis. Segundo o princípio da inclusão-exclusão, a união dos eventos é igual à soma da sua probabilidade menos a sua interceção, como se pode verificar na fórmula seguinte, num caso em que existem dois modos de falha:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (16)$$

Caso haja mais que dois modos de falha, é necessário subtrair a interceção entre cada par de eventos e somar a interceção de todos eles.

$$\begin{aligned} P(A \cup B \cup C) = & P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) \\ & - P(B \cap C) + P(A \cap B \cap C) \end{aligned} \quad (17)$$

Desta forma, a probabilidade de falha geral do ativo nunca passa dos 100% e é sempre no mínimo igual à probabilidade mais elevada dos modos de falha.

Voltando ao caso do disjuntor, os fatores identificados para cada modo de falha são apresentados na tabela 6 e os pesos atribuídos a cada um na tabela 7.

Modo de Falha	Características Técnicas	Condição Técnica (HI Modo Falha)	Factores Externos - I	Factores Externos - II	Condições operação
Perda Isolamento Interno	Idade	Índice de Saúde Isolamento Interno	Nível Cerâmico	Risco Incêndio (45%) Risco Eventos Atmosféricos Extremos (45%) Risco Sísmico (10%)	-
Perda Isolamento Externo	Idade	Índice de Saúde Isolamento Externo	Nível Cerâmico (45%) Poluição ou Contaminação (45%) Avifauna (10%)	Risco Incêndio (30%) Risco Inundações (20%) Risco Eventos Atmosféricos Extremos (20%) Risco Sísmico (30%)	-
Falha Operação ligar	Idade	Índice de Saúde Comando	Nível Cerâmico (30%) Poluição ou Contaminação (70%)	Risco Incêndio (30%) Risco Inundações (20%) Risco Eventos Atmosféricos Extremos (20%) Risco Sísmico (30%)	Tempo sem manobra
Falha Operação desligar	Idade	Índice de Saúde Comando	Nível Cerâmico (30%) Poluição ou Contaminação (70%)	Risco Incêndio (30%) Risco Inundações (20%) Risco Eventos Atmosféricos Extremos (20%) Risco Sísmico (30%)	Tempo sem manobra

Tabela 6 - Fatores impactantes nos modos de falha de um Disjuntor

Modo de Falha	Características Técnicas	Condição Técnica (HI Modo Falha)	Factores Externos - I	Factores Externos - II	Condições operação
Perda Isolamento Interno	10.00%	80.00%	7.50%	2.50%	-
Perda Isolamento Externo	10.00%	80.00%	7.50%	2.50%	-
Falha Operação ligar	10.00%	70.00%	7.50%	2.50%	10.00%
Falha Operação desligar	10.00%	70.00%	7.50%	2.50%	10.00%

Tabela 7 - Peso atribuído aos fatores impactantes nos modos de falha de um Disjuntor

Como se pode verificar, é dado um grande relevo ao Índice de Saúde. Teoricamente, é expectável que um disjuntor em condição crítica tenha uma maior probabilidade de falha.

A valorização da idade do disjuntor é feita através de uma função. Como se pode ver na tabela 8, são definidas probabilidades de falha para pontos específicos do ciclo de vida do disjuntor, sendo depois feita uma regressão exponencial para gerar a função, como demonstrado na figura 9.

Idade	PF
0	0.025
10	0.1
15	0.15
20	0.2
25	0.32
30	0.5
40	1

Tabela 8 - Relação entre a probabilidade de falha de um disjuntor e a sua idade

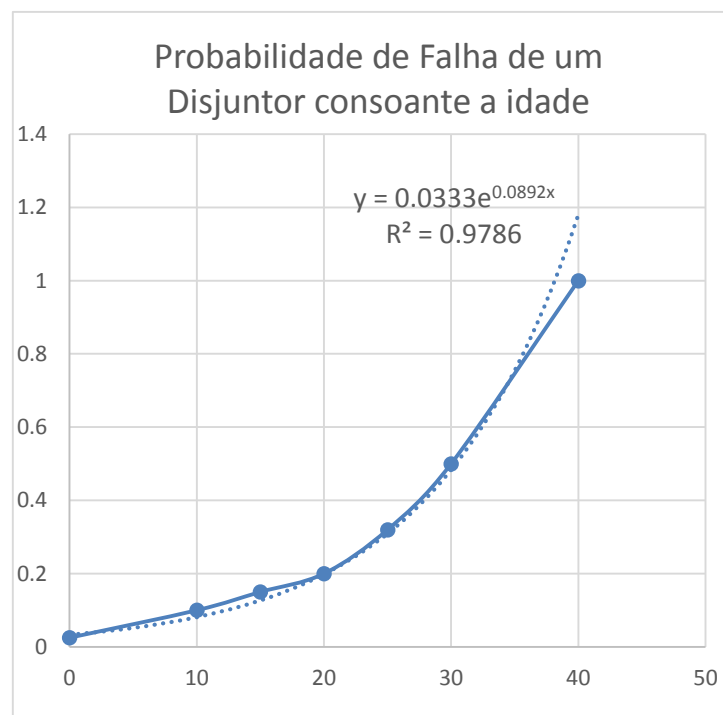


Figura 9 - Função da probabilidade de falha de um Disjuntor consoante a idade

Na tabela seguinte são apresentados os níveis considerados para contabilizar os fatores externos.

Fator Externo	Nível 0	Nível 1	Nível 2	Nível 3	Crítico
Poluição	Sem poluição	Poluição Industrial	Poluição Marítima - 1 km ≤ dist. mar ≤ 5 km	Poluição Marítima - < 1 km do mar	Poluição Marítima + Poluição Industrial
Nível Cerâmico	< 12	12-15	15-18	18-21	21-24
Avifauna	Sem Aves	Outras aves	Aves pequeno porte	Cegonhas	Aves pequeno porte + Cegonhas
Incêndio (zonas Geográficas)	Zona sem risco de incêndio	Zona de baixo risco	Zona de risco moderado	Zona elevado risco de incêndio	Zona de muito elevado risco de incêndio
Risco Sísmico	Muito Baixo	Baixo	Médio	Alto	Muito Alto
Inundação	Muito Baixo	Baixo	Médio	Alto	Muito Alto
Eventos Atmosféricos extremos	Muito baixa incidência	Baixa incidência	Incidência moderada	Alta Incidência	Muito Alta Incidência
Exposição a poluição / contaminação / climatização inadequada	Sem exposição	Exposição < 1 ano	Exposição: 1 a 2 anos	Exposição: 2 a 5 anos	Exposição > 5 anos
Tempo sem manobrar	Até 6 meses	6 meses a 1 anos	1 a 2 anos	2 a 5 anos	Mais de 5 anos

Tabela 9 - Fatores externos impactantes num disjuntor

3.2.3 Tempo de Vida Restante

O tempo de vida restante do ativo indica o tempo expectável em que este continuará em funcionamento, cumprindo a sua função. Este indicador é calculado de duas formas diferentes.

É feita uma estimativa em função da obsolescência do ativo, considerando cenários em que o ativo se pode tornar obsoleto por falta de componentes de reserva, *know-how* e assistência ou ser ultrapassado por outra tecnologia mais rentável. Além disso, podem haver alterações das necessidades da empresa, por exemplo devido a novas legislações, que façam com que o ativo deixe de cumprir todos os requerimentos.

O outro método é uma abordagem estatística complementada com parâmetros que tenham uma correlação comprovada com o tempo de vida do ativo. É definido um tempo de vida máximo para a tecnologia do ativo à qual se subtrai a idade deste. São consideradas as condições de operação do ativo, como a temperatura ou exposição a contaminantes, e parâmetros indicadores do seu estado de degradação para reduzir a estimativa, encurtando o tempo restante de vida para refletir as pressões a que o ativo está ou esteve sujeito. O resultado final é o tempo de vida restante mais curto de entre as duas estimativas.

No caso dos disjuntores foram definidos tempos de vida máxima para cada tecnologia, como apresentado na tabela 10. Para refletir a degradação sofrida pelo disjuntor são considerados três fatores, a Resistência de Contacto, a Corrente Cortada e o número de anos que o disjuntor esteve exposto a poluição, contaminação ou climatização deficiente.

Vácuo	SF6	Óleo
40 Anos	35 Anos	40 Anos


Tabela 10 - Tempo de vida máxima por tecnologia de disjuntor

3.2.4 Risco

O principal objetivo destes algoritmos é estimar a condição dos ativos de forma a identificar necessidades e oportunidades. No entanto, é necessário priorizar as intervenções identificadas como necessárias para obter o melhor resultado possível, diminuindo os custos das intervenções e evitando potenciais prejuízos no futuro. Através da associação das estimativas da probabilidade de falha do ativo com as consequências de este falhar é possível aferir a sua criticidade. Na EDP Distribuição esta associação é feita através da Matriz de Risco apresentada na figura 10.

Caso a caracterização do estado do ativo seja fidedigna então esta contribui de forma significativa para o processo de decisão, permitindo a priorização das intervenções de forma objetiva.

Na figura 11 é resumida a aplicabilidade dos indicadores anteriormente apresentados para as decisões de manutenção dos ativos.



Matriz de Risco - 2010

EDP Distribuição - Gestão Riscos Empresariais

Impactos											Frequência				
Valores do Negócio	Sustentabilidade		Reputação	Qualidade de Serviço	Económicos	Período Médio entre Ocorrências (em anos)									
Indicadores	Segurança de Pessoas	Ambiente	Repercussão nos Média e População	TIEPI MT Interno (min) (base EDP)	Resultados (k€)	muito elevada	elevada	média	baixa	muito baixa					
						(f ≤ 0,5)	(1 ≤ f < 0,5)	(2 ≤ P < 1)	(5 ≤ f < 2)	(5 ≤ f)					
						5	4	3	2	1					
Nível de Severidade	5 muito crítico	Pode causar morte ou incapacidade permanente e grave de pessoas	Pode causar danos significativos ao ambiente que se manifestam durante um prazo superior a 5 anos	- Internacional - ou nacional com potencial de repetição nos médias	TIEPI MT ≥ 3	R ≥ 4.500	I1	I2	I4	M5	M1				
	4 crítico	Pode necessitar de internamento hospitalar	Pode causar danos significativos ao ambiente que se manifestam durante um prazo até 5 anos	- Nacional - ou regional com potencial de repetição nos médias	3 > TIEPI MT ≥ 1,33	4.500 > R ≥ 2.000	I3	I5	M6	M2	A10				
	3 significado alto	Pode necessitar tratamento médico	Pode causar danos pequenos ao ambiente que se manifestam durante mais de 5 anos	- Regional - ou local com potencial de repetição nos médias	1,33 > TIEPI MT ≥ 0,5	2.000 > R ≥ 750	I6	M7	M3	A9	A6				
	2 significado médio	Pode necessitar de Primeiros Socorros	Pode causar danos pequenos ao ambiente que se manifestam durante um prazo até 5 anos	Notícia local	0,5 > TIEPI MT ≥ 0,17	750 > R ≥ 250	M8	M4	A8	A5	A3				
	1 significado baixo	Sem impacto	Sem impacto relevante	Sem impacto externo ao Grupo	0,17 > TIEPI MT	250 > R	A11	A7	A4	A2	A1				

Figura 10 - Matriz de Risco EDP Distribuição

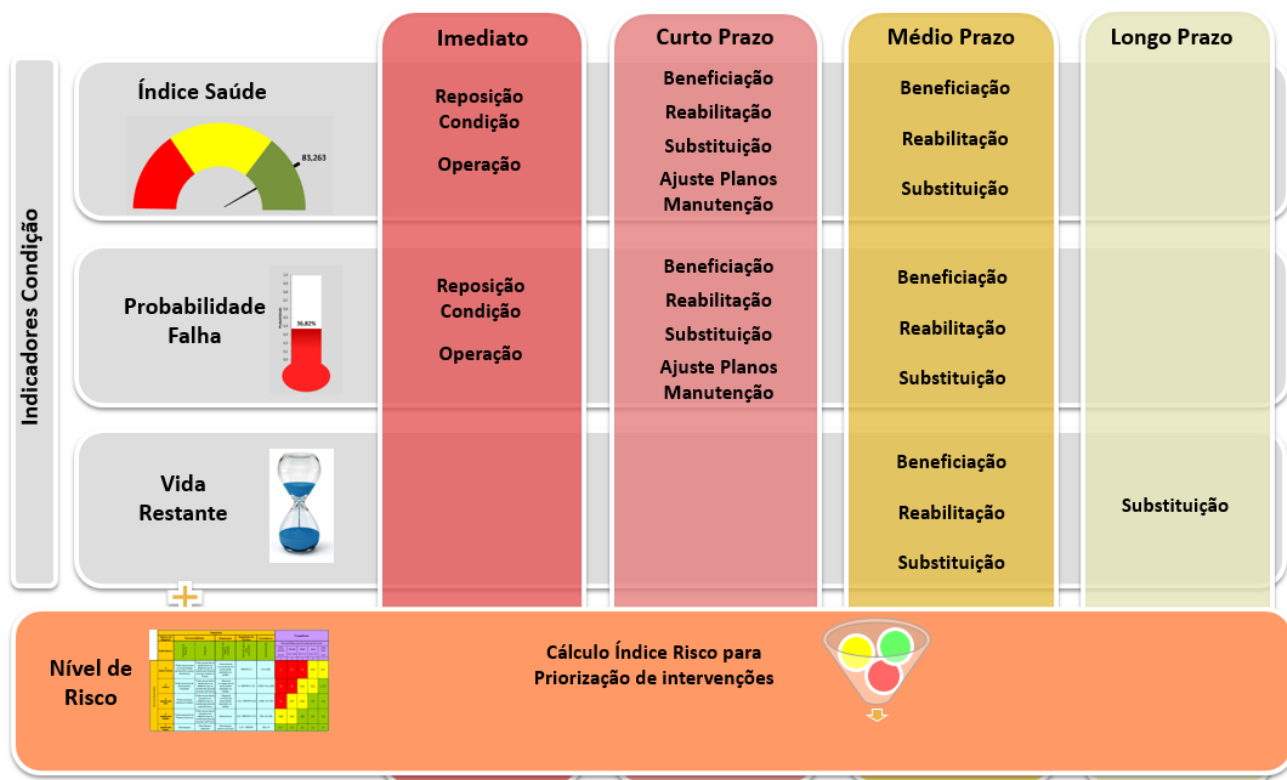


Figura 11 - Aplicabilidade dos indicadores de condição

3.3 Análise de ensaios em disjuntores

Um dos principais métodos de recolha de informações de um disjuntor é a realização de ensaios que permitem aferir alguns parâmetros críticos para a sua *performance*. Nesta secção é apresentada uma análise realizada num conjunto de ensaios efetuados em vários disjuntores com o intuito de encontrar valores de referência para os parâmetros estudados.

Como foi referido anteriormente, a principal função de um disjuntor é a de interromper a corrente quando necessário. Uma das situações em isto pode acontecer é, por exemplo, quando um ramo de uma árvore toca numa linha provocando um curto-circuito. O disjuntor deve atuar rapidamente, fazendo um corte de tensão na rede elétrica em que está inserido. Desta forma evita-se que a anomalia se propague e provoque danos noutros equipamentos.

Os valores em estudo nos ensaios recolhidos são os seguintes:

- Tempo de Abertura;
- Tempo de Fecho;
- Resistência de Contacto.

Os disjuntores em análise são trifásicos o que significa que têm três polos por onde passa a corrente, sendo todos acionados quando é necessário fazer uma manobra. Os valores são medidos para cada um dos polos do disjuntor sendo considerada a média dos resultados obtidos para simplificar o processo de análise.

O tempo de abertura de um disjuntor é o intervalo de tempo entre o disjuntor receber um comando de abertura e o momento em que este termina a manobra de abertura dos polos, interrompendo a corrente. Este valor é medido em milissegundos e denota a rapidez de resposta do disjuntor. O tempo de abertura não pode ser elevado para que o disjuntor atue a tempo de impedir a propagação da falha na rede elétrica.

O tempo de fecho é semelhante ao tempo de abertura mas aplicado à manobra de fecho do disjuntor. Depois de a falha ser resolvida os polos devem retornar à posição inicial de forma a restabelecer a corrente elétrica.

A interrupção da corrente na manobra de abertura dos polos gera um arco de eletricidade que provoca a erosão dos seus pontos de contacto. A resistência de contacto dos polos do disjuntor é medida para aferir a sua capacidade de extinção dos arcos gerados. Tal como nos tempos de abertura e de fecho quanto menor este valor melhor o estado do disjuntor. Um disjuntor com resistência de contacto elevada pode sofrer grandes subidas de temperatura quando ocorre um arco e comprometer a segurança dos seus componentes.

Foram agregados os dados de vários ensaios efetuados em disjuntores com o intuito de encontrar os valores de referência para os parâmetros referidos. Estes valores são utilizados no cálculo do Índice de Saúde de um disjuntor, sendo este estudo uma importante contribuição para a correta caracterização do estado de degradação deste tipo de ativo.

Ao todo foram analisados 1195 ensaios. Os disjuntores foram divididos por meio de interrupção: Ar, Óleo, SF6 ou Vácuo. Devido a alguns erros na designação do modelo dos disjuntores foi necessário limpar e formatar os dados de forma a garantir o máximo de associações possíveis. No gráfico da figura 12 é apresentada a distribuição dos ensaios à resistência de contacto por meio de interrupção do disjuntor. O número reduzido de ensaios em disjuntores a ar comprimido deve-se a estes se terem tornado obsoletos e estarem a ser substituídos no tipo de subestação em estudo.

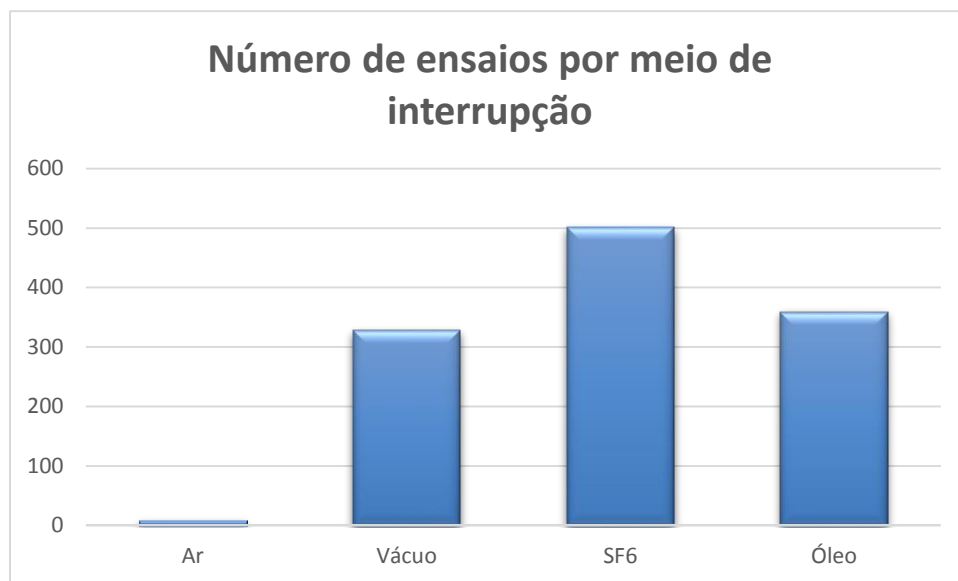


Figura 12 - Número de ensaios por meio de interrupção de disjuntor

O tipo de gráfico aplicado para facilitar a análise dos ensaios foi o *boxplot*. Este tipo de gráfico permite comparar a variação de uma variável entre diferentes grupos de dados. São calculados a mediana e os quartis dos dados e é desenhada uma caixa que vai do primeiro quartil ao terceiro quartil, ficando de fora desta os 25% valores menores e os 25% valores maiores. A partir dos extremos da caixa são desenhadas duas linhas verticais, uma superior e outra inferior, designadas como *whiskers*, que vão da caixa até ao valor mais próximo de um limite predefinido. O limite utilizado é a região interquartil, que é a distância entre o primeiro e o terceiro quartil, multiplicada por 1,5. Os valores que fiquem de fora destes limites são considerados *outliers*, não influenciando o tamanho da caixa. Como o objetivo da análise é encontrar os valores de referência a que se espera que os disjuntores se aproximem é importante excluir estes valores fora do normal que podem ter surgido devido a erros de medição ou de registo.

O gráfico da figura 13 representa a média da resistência de contacto nos três polos de um disjuntor, sendo os ensaios separados pelo meio de interrupção deste. Cada ponto representa um ensaio e linha horizontal dentro da caixa é a mediana.

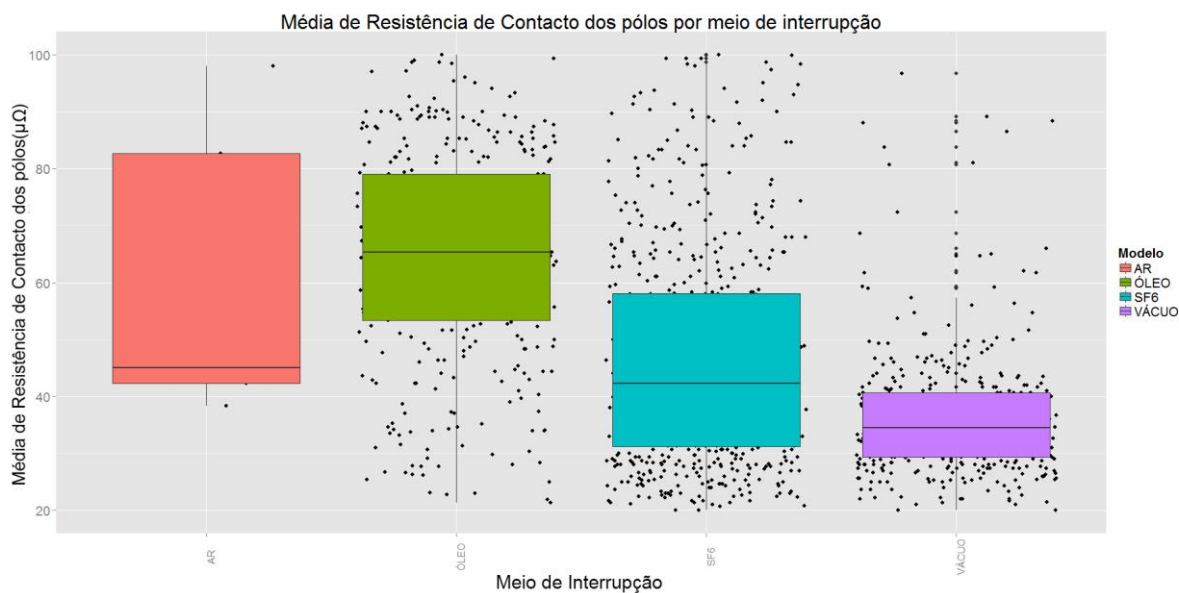


Figura 13 - Resistência de Contacto dos polos por meio de interrupção

Como seria de esperar devido ao número reduzido de ensaios em disjuntor a ar comprimido este tipo de gráfico não se aplica bem a esse grupo. No entanto, esta tecnologia não é tida em conta no algoritmo de avaliação do Índice de Saúde para o disjuntor. É possível identificar diferenças entre os ensaios nas restantes tecnologias, sendo que este estudo serviu de base para calibrar os níveis do parâmetro relativo à resistência de contacto em cada tecnologia no algoritmo do Índice de Saúde.

Caso se pretenda distinguir os disjuntores também por nível de tensão (Média ou alta) e por modelo este tipo de gráfico também tem utilidade, permitindo identificar as tendências de cada modelo e associações entre modelos com valores semelhantes. No gráfico da figura 14 estão representados os ensaios efetuados em disjuntores de Média Tensão cujo meio de interrupção é a óleo. Os restantes gráficos produzidos, para a Resistência de Contacto, Tempo de Abertura e Tempo de Fecho podem ser consulados em anexo.

Este tipo de análise estatística permite adquirir conhecimento com valor para a manutenção de ativos, potencialmente reduzindo a subjetividade no seu diagnóstico e refinando a caracterização do seu estado.

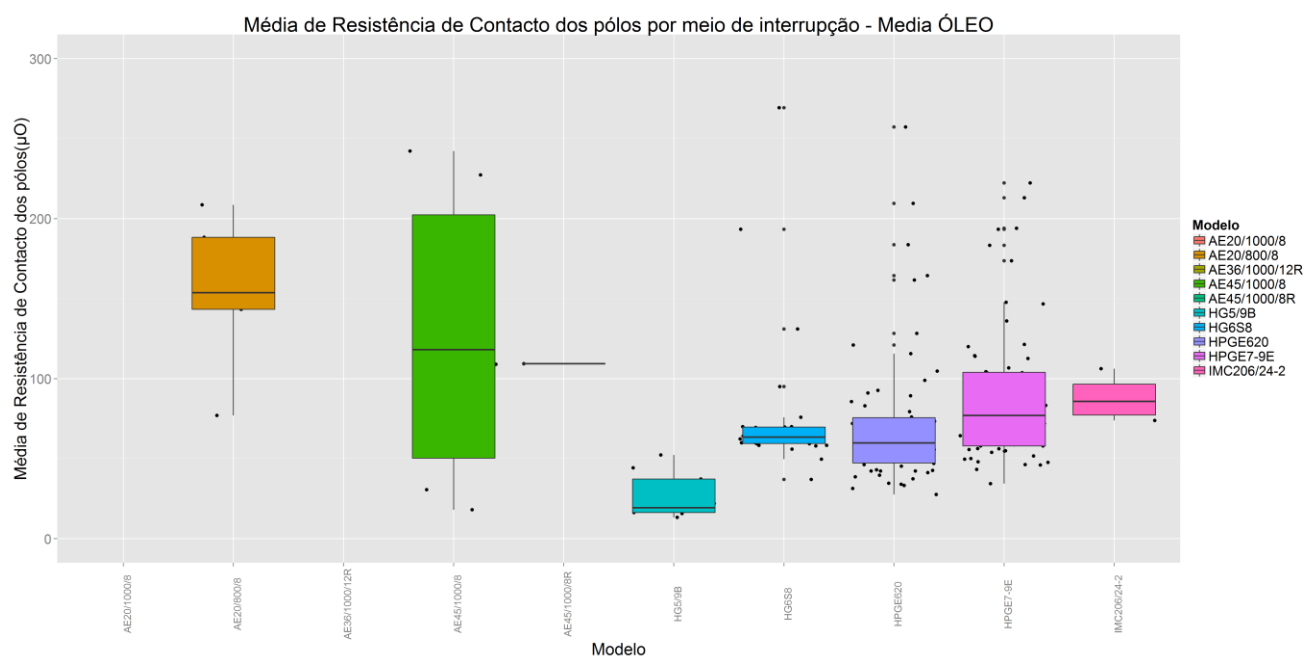


Figura 14 - Resistência de Contacto de disjuntores a Óleo e de Média Tensão

Capítulo 4

Aplicações de Data Mining para Manutenção Preditiva de uma subestação

Uma subestação elétrica é um sistema muito complexo, composto por vários ativos com características diferentes, tal como já foi referido em capítulos anteriores. A evolução dos sistemas de monitorização permite o registo de vários parâmetros relevantes para o processo de decisão na manutenção destes ativos. No entanto, a elevada complexidade dos ativos em questão implica uma grande quantidade de dados, o que dificulta a tarefa de avaliar manualmente a sua condição operacional de forma fundamentada e objetiva. A área de *Data Mining* surge como resposta a esta necessidade de identificar padrões nos dados e extrair conhecimento valioso.

Neste capítulo são estudadas algumas abordagens de *Data Mining* para auxiliar o processo de manutenção preditiva de uma subestação elétrica tendo em vista a sua potencial replicação no contexto da EDP Distribuição. São apresentados métodos que possam contribuir para a estimação dos indicadores abordados no capítulo anterior.

É identificada uma abordagem que consiste em aplicar métodos de *Survival Analysis* para estimar a probabilidade de Falha relativa à idade de um disjuntor com o objetivo de melhorar o algoritmo existente. Este caso de estudo é apresentado e discutido.

4.1 Revisão do Estado da Arte

Nesta secção é feita uma revisão do Estado da Arte de aplicações de métodos de Data Mining equivalentes aos algoritmos desenvolvidos na EDP para estimação do Índice de Saúde, da Probabilidade de Falha e do tempo de vida restante. Esta análise centra-se no Transformador de Potência uma vez que este é o ativo mais estudado e o mais valioso de uma subestação elétrica.

4.1.1 Índice de Saúde

Na avaliação da condição de funcionamento de um Transformador de Potência através do Índice de Saúde foram utilizados métodos diversificados, incluindo abordagens no âmbito da Inteligência Artificial. Nesta secção serão apresentadas aplicações de *Fuzzy Logic*, Redes Neurais e *Support Vector Machine* para este problema.

4.1.1.1 Fuzzy Logic

Os métodos de *Fuzzy Logic*, ou Lógica Difusa, podem ser utilizados para codificar o conhecimento empírico e experiência adquirida na manutenção de um transformador e na avaliação da sua condição. Os modelos criados são capazes de suportar valores não exatos ou variáveis qualitativas e simular processos de raciocínio. Assim, é possível traduzir em termos matemáticos conhecimento impreciso ou uma avaliação subjetiva, permitindo atribuir um grau de pertinência de um elemento a um conjunto difuso. O grau de pertinência representa o quanto um elemento pertence a um conjunto. Esta característica é importante porque algumas das técnicas utilizadas na indústria para avaliar a condição de transformadores, como inspeções visuais, podem ter como resultado uma categoria e não um valor exato, por exemplo “Mau” ou “Muito Bom”.

Em Arshad et al. (2014) é proposto um modelo de *Fuzzy Logic* que integra vários parâmetros de diagnóstico e constrói um conjunto de regras e conclusões para simular o processo de decisão de um engenheiro experiente. As regras são utilizadas para classificar um elemento num conjunto e quantificar as criticidades do transformador e são baseadas em limites definidos em *standards* reconhecidos internacionalmente como o CIGRE e o IEC.

Regras difusas “IF-THEN” foram criadas para representar o conhecimento extraído do seguinte processo:

- Um conjunto de transformadores em diferentes condições foi monitorizado durante 24 meses;
- Foram realizados vários testes de diagnóstico e os resultados foram interpretados com o apoio de engenheiros experientes na manutenção de transformadores;
- Foi avaliada a relação entre as características de um transformador e a sua resposta a pressões externas para criar regras difusas e ajustar os graus de pertinência de forma a representar os seus valores de criticidade.

As criticidades de um transformador que foram identificadas e os seus indicadores são apresentados na tabela 11. As criticidades são agrupadas em sequências com base nas dependências existentes entre si. Na figura 15 é apresentado um exemplo para o caso da criticidade térmica (*Thermal Criticality*). O resultado final do modelo é a criticidade geral do transformador que resulta da integração de todas as criticidades identificadas. O significado deste valor é semelhante ao de um Índice de Saúde, é um valor numérico que pode ir de 0 a 1 em que 0 representa um transformador em excelentes condições e 1 representa um transformador muito degradado.

S No.	Criticalities	Indicators
1	Thermal	
	Oil	DGA-Ethylene and Ethane
	Paper	DGA-Carbon Monoxide and Carbon Dioxide
2	Electrical	
	Partial Discharge	DGA-Hydrogen and Methane
	Arcing	DGA-Hydrogen and Acetylene
3	Active	
	Prognostic	DGA-Rate of Gas Evolution and HST
	Operational	Loading and Environmental effects
4	Winding Integrity	
	Electrical	Partial Discharge and Leakage Reactance
	Mechanical	FRA and DDF trend
5	Oil Quality	
	Electrical	Dielectric Breakdown Voltage and PF
	Physical	IFT and UV Spectrophotometry
6	Insulation	
	Deterioration	Moisture (RVM) and HST

Tabela 11 - Criticidades de um Transformador (Arshad, Islam, and Khaliq 2014)

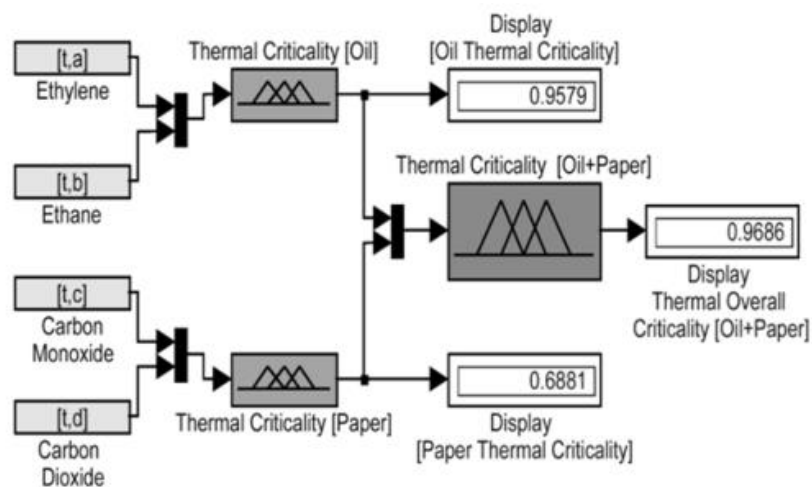


Figura 15 - Modelo difuso para criticidade térmica (Arshad, Islam, and Khaliq 2014)

Os dados recolhidos através da monitorização dos transformadores foram divididos em conjunto de treino, utilizado para construir o modelo, e em conjunto de teste, para verificar a sensibilidade do modelo e a sua precisão. As criticidades identificadas estiveram de acordo com as falhas que ocorreram nos transformadores utilizados como teste.

Este método tem a desvantagem de ser dependente do conhecimento extraído, o que requer uma elevada especialização e habilidade de monitorizar os transformadores de forma a adquirir dados relevantes. É interessante a sua capacidade de identificar claramente os fatores que mais prejudicam a condição de um transformador devido a decompor o processo de avaliação em vários modelos, correspondentes a cada criticidade. Tendo em conta as criticidades mais elevadas podem ser realizados testes de diagnóstico mais específicos para identificar potenciais falhas num disjuntor.

4.1.1.2 Rede Neuronal

Uma Rede Neuronal é capaz de aprender diretamente dos dados disponíveis e ajustar as suas estimações a partir destes. Nos métodos de *Fuzzy Logic*, como o apresentado previamente, as regras de avaliação têm que ser definidas previamente, o que pode ser considerado uma limitação do método. As Redes Neurais apresentam uma opção muito interessante na abordagem do problema de diagnóstico do estado de um transformador, minimizando o aspeto subjetivo deste tipo de análise.

Mukti et al (2014) propõem uma abordagem baseada em Redes Neurais. Tendo como base um método de diagnóstico de transformadores proposto pelo Departamento do Interior dos Estados Unidos da América (Reclamation 2003), que identifica alguns fatores como parâmetros primários e outros como parâmetros secundários, o modelo proposto é composto por dois níveis de Redes Neurais. Os parâmetros utilizados e a estrutura do modelo podem ser vistos na seguinte figura.

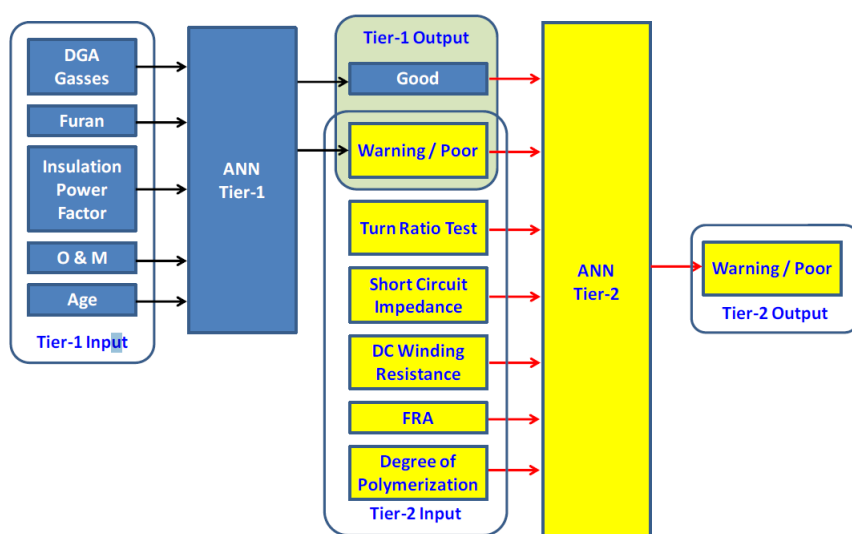


Figura 16 - Aplicação de Redes Neurais para avaliação de transformadores (Mukti, Pamuji, and Munir 2014)

Em cada nível é utilizada Rede Neuronal para resolver um problema de Classificação em que a variável objetivo é o estado do transformador, tendo como valores possíveis “Warning”, “Poor” e “Good”. A Rede Neuronal do segundo nível recebe a classificação resultante da primeira Rede Neuronal como variável de entrada. Esta abordagem foi desenvolvida de forma a ser mais robusta a casos em que o transformador tenha mais que um tipo de falha ao mesmo tempo.

Os dados utilizados consistem em aproximadamente 16 mil medidas de transformadores em funcionamento, tendo sido utilizadas 90% das amostras para treinar o modelo e os restantes 10% para o testar. Como este é um método de Aprendizagem Supervisionada, tem que ser atribuído um valor para o resultado esperado de cada amostra para servir de referência ao modelo. Este valor foi calculado manualmente de acordo com o método de diagnóstico que serviu de inspiração para esta abordagem.

O modelo obteve uma percentagem de acerto muito elevada, conseguindo 99.5% de acerto nas suas previsões. Foi utilizado o método de *10-fold Cross Validation* para testar a robustez do modelo, sendo que este demonstrou uma boa capacidade de generalização.

Outra abordagem a este problema através de Redes Neurais é apresentada em Abu-Elanien et al. (2011). São utilizados parâmetros diferentes e os valores de referência para o Índice de Saúde das amostras é providenciado por uma empresa especializada que utiliza um método de avaliação confidencial. Apenas uma Rede Neuronal é utilizada e tem como base as propriedades do óleo isolante do transformador. Quando este sofre pressões são libertados gases no óleo que permitem identificar degradações ocorridas. O modelo está representado na figura 17.

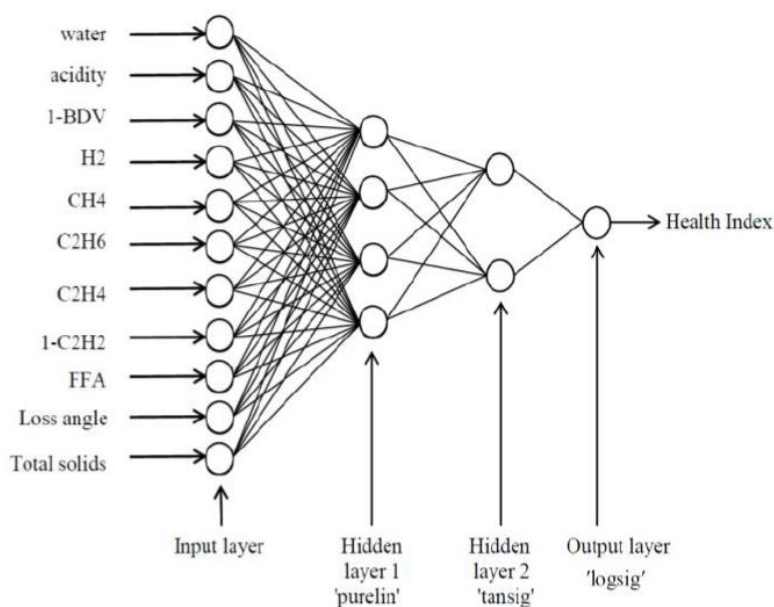


Figura 17 - Rede Neuronal proposta por Abu-Elanien et al (2001)

Foram utilizadas as amostras de 59 transformadores para treinar o modelo e este foi testado com amostras de outros 29 transformadores. O resultado do modelo é um Índice de Saúde de 0 a 1 em que quanto menor o valor melhor o estado do transformador. Este modelo também obteve uma percentagem de acerto muito elevada de 96.5%.

Este estudo demonstra o potencial das aplicações de Redes Neurais para estimar a condição de um transformador, nomeadamente a sua capacidade de, depois de o modelo estar construído, este ser capaz de receber diferentes tipos de parâmetro e calcular automaticamente o Índice de Saúde de um transformador sem ser necessário conhecimento especializado. No entanto, estas abordagens estão intimamente dependentes de uma estimação feita previamente através de outro método. Há uma relação direta entre a qualidade das estimações do modelo e as estimações feitas pelo modelo que lhe serviu de referência, ou seja, se o método de referência não avaliar de forma correta o estado do transformador então o modelo desenvolvido também não será capaz de o fazer. Os métodos apresentados têm o mérito de automatizar o cálculo do Índice de Saúde e de atingir resultados semelhantes, evitando que seja necessário contratar uma empresa especializada para o fazer novamente no futuro. Na sua utilização deve ser tida em conta, no entanto, a seguinte condição: é necessário uma forma de avaliar o acerto das previsões para que um método de Aprendizagem Supervisionada como as Redes Neurais seja capaz de contribuir com um resultado útil e que realmente traga conhecimento novo.

4.1.1.3 Support Vector Machine

Em Ashkezari et al (2014), são utilizadas técnicas de seleção de variáveis para encontrar o subconjunto mais informativo de um conjunto de variáveis resultantes de testes ao óleo isolante de um transformador. Pretende-se descobrir as variáveis com mais correlação com o Índice de Saúde e, ao mesmo tempo, reduzir a redundância no cálculo deste indicador. Os subconjuntos são utilizados para treinar um modelo de *Support Vector Machine* (SVM) e determinar o Índice de Saúde do sistema de isolamento de um transformador. Especificamente, as técnicas utilizadas foram a análise de correlação e *Minimum Redundancy Maximum Relevance* (mRMR).

Foram tidas em conta doze características do óleo relativas à sua acidez, humidade, gases dissolvidos, propriedades elétricas, etc. Foi construído um conjunto de dados com estas amostras e o Índice de Saúde respetivo, calculado através de um método proposto pelos mesmos autores baseado em conhecimento especializado e em *standards* da indústria (Ashkezari et al. 2013). O estado do isolamento do transformador pode ser classificado por quatro classes: “Excelente”, “Bom”, “Razoável” e “Mau”.

O procedimento adotado para criar o modelo foi o seguinte:

- Obtenção de um subconjunto de variáveis através de técnicas de seleção de variáveis;
- Normalização das amostras e divisão em conjuntos de treino e de teste;

- Utilização de *10-fold Cross Validation* para otimizar os parâmetros do modelo SVM;
- Construção do modelo com o conjunto de treino;
- Cálculo da percentagem acerto das previsões do modelo no conjunto de teste;
- Repetição do processo vinte vezes para obter a percentagem de acerto média do modelo.

Através da análise de correlação, as variáveis com uma elevada correlação são agrupados no valor comum, tendo sido obtidas sete variáveis a partir das doze originais. Na tabela 12 são comparados dois modelos SVM utilizando como conjunto de teste 350 amostras de óleo, sendo que um tem como base todas as variáveis e outro utiliza o subconjunto de sete variáveis. É analisado o tempo de computação de ambos os modelos e a percentagem de acerto média obtida no conjunto de teste e a percentagem de acerto por classe. É possível verificar que o modelo baseado no subconjunto de variáveis obteve uma percentagem de acerto ligeiramente maior com um tempo de computação menor, devido à redução da dimensionalidade dos dados.

Número de variáveis	Tempo de computação (segundos)	Percentagem de acerto médio	Percentagem de acerto por classe			
			Excelente	Bom	Razoável	Mau
12	753	86	86	82	83	94
7	602	87	87	84	84	94

Tabela 12 - Comparação de SVM com e sem análise de correlação (Ashkezari et al. 2014)

Com a técnica mRMR são selecionados subconjuntos com um número de variáveis em crescendo com o objetivo de minimizar a redundância e maximizar a relevância as variáveis de entrada no modelo. A tabela 13 apresenta uma comparação da percentagem de acerto dos vários modelos criados. Os modelos com menos de onze variáveis obtém melhor percentagem de acerto que o modelo original.

Número de variáveis	Percentagem de acerto médio	Percentagem de acerto por classe			
		Excelente	Bom	Razoável	Mau
7	87	83	84	85	93
8	87	84	84	85	95
9	87	85	82	88	94
10	87	88	84	85	93
11	86	86	83	83	93
12	86	86	82	83	94

Tabela 13 - Comparação de modelos SVM com diferente número de variáveis (Ashkezari et al. 2014)

Ambas as técnicas de seleção de variáveis obtiveram resultantes interessantes pois permitiram ao modelo SVM ser mais eficiente, reduzindo a sua complexidade computacional sem afetar a percentagem de acerto das suas previsões, conseguindo até aumentá-la ligeiramente. As técnicas de seleção de variáveis devem ser consideradas quando se lida com dados com elevada dimensionalidade como neste caso, visto que são gerados muitos dados a partir da monitorização de um equipamento complexo como um Transformador de Potência.

Este estudo demonstra que o método de *Support Vector Machine*, tal como as Redes Neurais, consegue estimar de forma precisa o Índice de Saúde de um transformador e é capaz de lidar com dados muito heterogêneos. No entanto, o mesmo problema verificado nas Redes Neurais de depender de um Índice de Saúde de referência também se aplica neste caso e a conclusão principal é que é necessário desenvolver um método para verificar as previsões obtidas por um método de Aprendizagem Supervisionada para que o seu resultado não seja apenas uma estimativa de uma estimativa.

4.1.2 Probabilidade de Falha

Algumas abordagens baseadas em Inteligência Artificial foram já utilizadas com sucesso para melhorar a deteção de falhas em Transformadores de Potência. O método mais comum é o uso de algoritmos de Aprendizagem Supervisionada para aumentar a precisão das técnicas de *Diagnostic Gas Analysis* (DGA), que se baseiam nos gases dissolvidos no óleo isolante do transformador. As técnicas de DGA são capazes de identificar diferentes tipos de falha, de natureza térmica ou elétrica.

Este é um problema de Classificação em que o modelo tem que determinar relações entre as concentrações dos gases no óleo e os possíveis estados de falha ou funcionamento normal. Para que seja possível utilizar um método de Aprendizagem Supervisionada para esta abordagem é necessário reunir um conjunto de amostras de óleo e o estado do transformador no momento de recolha, obtido através de diagnósticos mais complexos ou conhecido posteriormente, caso ocorra uma falha nesse transformador. A vantagem de utilizar métodos baseados em técnicas de DGA é que com os avanços tecnológicos recentes é possível utilizar sensores para monitorizar em tempo real os níveis dos gases no óleo.

Um dos métodos mais utilizados para esta abordagem é a combinação de vários modelos *Support Vector Machine* em camadas (Bacha, Souahlia, and Gossa 2012, Lian and Ma 2013, Ganyun et al. 2005), consistindo numa classificação hierárquica. Na figura está representada a estrutura de um método deste tipo. O primeiro classificador distingue entre um estado de funcionamento normal e um estado de falha. Os classificadores subsequentes identificam o tipo de falha de entre seis possibilidades definidas pela norma internacional IEC 60599(60599 2007). Na tabela 14 são apresentados os tipos de falhas previstas e o seu código.

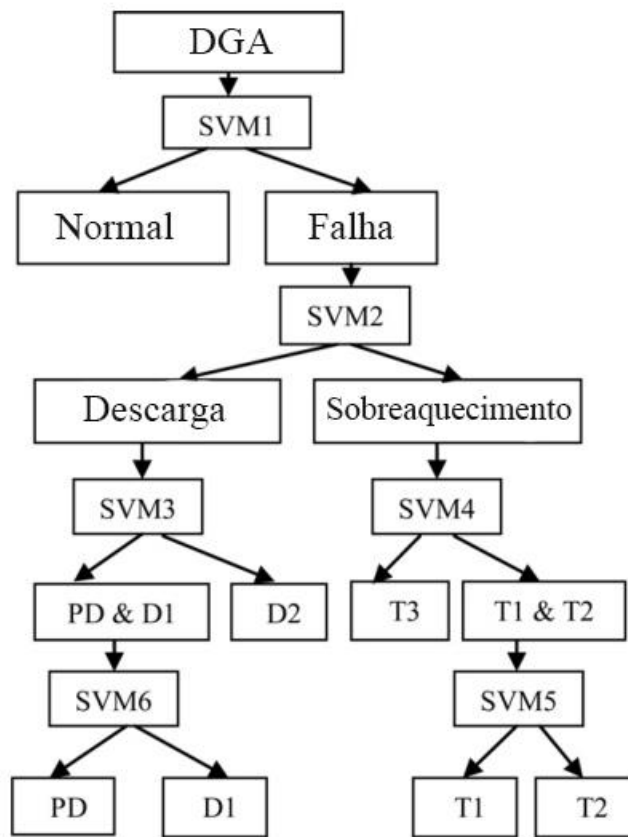


Figura 18 - Multi-Layer SVM para detecção de falha (Bacha, Souahlia, and Gossa 2012)

Tipo de Falha	Código
Descarga parcial	PD
Descarga de baixa energia	D1
Descarga de alta energia	D2
Falha Térmica, $T < 300\text{ }^{\circ}\text{C}$	T1
Falha Térmica, $300\text{ }^{\circ}\text{C} < T < 700\text{ }^{\circ}\text{C}$	T2
Falha Térmica, $T > 700\text{ }^{\circ}\text{C}$	T3

Tabela 14 - Código dos tipos de falha por DGA (Bacha, Souahlia, and Gossa 2012)

Métodos de Redes Neurais e *Fuzzy Logic* também já foram utilizados com sucesso para esta abordagem (XiaoDong and Hongzhi 2008, Liu et al. 2008, Nemeth et al. 2010). Nos casos de estudo analisados onde foram comparados diferentes modelos no mesmo conjunto de dados, o modelo de SVM, talvez devido à sua maior robustez a situações em que estejam disponíveis poucos dados de treino, obteve uma percentagem ligeiramente superior (Bacha, Souahlia, and Gossa 2012, Ganyun et al. 2005). A tabela seguinte apresenta uma comparação feita em Bacha, Souahlia e Gossas (2012) das percentagens de falso alarme, de não deteção de falha e de precisão do diagnóstico num caso de estudo um modelo de *Fuzzy Logic*, duas Redes Neurais e um modelo de SVM.

Método	Falso alarme (%)	Não-deteção (%)	Precisão do diagnóstico
<i>Fuzzy Logic</i>	6.7	6.7	86.7
MLP	0	20	80
RBF	0	13.3	86.7
SVM	0	10	90

Tabela 15 - Comparação de métodos de IA para deteção de falhas (Bacha, Souahlia, and Gossa 2012)

A aplicação de técnicas de seleção de variáveis para reduzir a redundância também já foi estudada tendo-se traduzido em melhorias na precisão dos modelos (Han, Hou-jun, and xiucheng 2011, XiaoDong and Li 2009, Lee et al. 2007).

Estes modelos conseguem, de facto, produzir resultados mais precisos que os métodos empíricos de DGA, especialmente em casos em que a concentração dos gases esteja perto dos limites definidos nesses métodos. No entanto, só é possível construir estes modelos se existirem amostras de óleo de transformadores quando ocorreram falhas de forma a reunir um conjunto de dados que permitam treinar os modelos. As amostras disponíveis podem ser divididas em conjuntos de treino e de teste para validar as previsões de um modelo, verificando se este deteta corretamente quando um ativo está em risco de falhar. Se houver outros parâmetros que tenham correlação com os estados de falha então estes modelos de Aprendizagem Supervisionada podem ser estendidos para os ter em conta e possivelmente alcançar resultados mais precisos e fundamentados.

4.1.3 Tempo de vida restante

O tempo de vida restante de um Transformador de Potência é o número de anos durante os quais se espera que este seja capaz de continuar a cumprir a sua função. Existem algumas propriedades do transformador com uma correlação comprovada empiricamente com o seu tempo

de vida restante, como o Grau de Polimerização do seu papel isolante ou a temperatura no ponto mais quente. No entanto, a medição destes fatores pode ser pouco precisa e exigir a interrupção de serviço do transformador.

Se for conhecida a idade em que vários transformadores com características semelhantes tenham tido uma falha geral, tendo que ser retirados, então esse conhecimento pode ser extrapolado para estimar várias métricas, como o tempo médio de vida ou a percentagem de transformadores que se espera que falhem antes de um determinado momento. Uma das restrições comuns neste tipo de análise é que uma grande parte da população conhecida esteja ainda em funcionamento. Isto é especialmente frequente no caso dos transformadores pois estes têm um tempo de vida longo, de trinta a quarenta anos. Além disso, fatores como as condições de operação ou o meio em que está inserido podem influenciar a longevidade de um transformador o que pode obrigar a dividir os disjuntores conhecidos em grupos, diminuindo o tamanho da amostra.

Para estimar o tempo de vida restante dos transformadores podem ser utilizados métodos de *Survival Analysis*, que tiram partido tanto dos transformadores abatidos, considerados indivíduos censurados, como dos que estão ainda no ativo. A abordagem deste tipo mais recorrente em transformadores é a aplicação de Métodos Paramétricos, tentando modelar o seu tempo de vida através de distribuições estatísticas. Nomeadamente, as distribuições *Weibull*, Exponencial e Log-normal já foram aplicadas com sucesso, permitindo descrever a distribuição do tempo de vida de um grupo de transformadores e prever com alguma segurança as suas falhas (Dan, Zhongdong, and Chengrong 2013, Maciejewski, Anders, and Endrenyi 2011, Badune, Vitolina, and Maskalonoks 2013, Lee et al. 2008). A distribuição Normal é menos indicada para este tipo de estudo porque prevê a existência de probabilidades para valores negativos e um tempo de vida é no mínimo igual a zero.

Além de relacionarem a idade com a probabilidade de sobrevivência, estes métodos também permitem calcular o tempo de vida médio de uma população com indivíduos censurados. Este valor pode ser comparado com os indivíduos com um tempo de vida totalmente conhecido através de técnicas como o Erro Quadrático Médio.

Em Zhou et al (2013) é analisado o efeito de utilizar os dados de sobrevivência dos transformadores ainda em funcionamento na precisão de modelos de estimação do tempo de vida. É provado que a precisão é aumentada dependendo do tamanho da amostra e da percentagem de indivíduos censurados na população, pelo que é sugerida a partilha de informação entre entidades que operem o mesmo tipo de equipamento e que o modelo seja atualizado quando são conhecidos novos dados de falha. É recomendado um número mínimo de vinte indivíduos não censurados para aplicar uma distribuição de *Weibull*.

Caso não seja possível ajustar uma distribuição estatística aos dados conhecidos podem ser utilizados métodos Não Paramétricos como *Kaplan-Meier* ou *Cox Proportional Hazards* para estimar a probabilidade de sobrevivência de uma população (Maciejewski, Anders, and Endrenyi 2011, Qing et al. 2011). Os modelos resultantes destes métodos não podem ser utilizados para fazer previsões fora dos limites dos dados utilizados para os treinar.

4.1.4 Conclusões

A estimação do Índice de Saúde, da probabilidade de Falha e do tempo de vida restante de um ativo proporciona ao gestor valores facilmente compreensíveis e ajudam a reduzir a subjetividade da sua análise. Foram estudados alguns métodos de *Data Mining* para a estimação destes indicadores analisando o seu interesse para a EDP Distribuição.

Relativamente ao Índice de Saúde foi demonstrado que alguns modelos de Aprendizagem Supervisionada são capazes de lidar com os diferentes tipos de parâmetros gerados por um ativo de uma subestação e estimar o seu estado de degradação. No entanto, a principal dificuldade deste tipo de abordagem é que é necessário haver uma estimação prévia que o modelo possa usar como referência. Assim sendo, e tendo em conta os objetivos da EDP Distribuição, os métodos apresentados para a estimação deste indicador não têm grande utilidade uma vez que não contribuem com conhecimento novo em relação ao algoritmo já existente. O resultado de um destes modelos seria uma estimação da estimação feita pelo algoritmo desenvolvido na EDP.

A aplicação com sucesso de métodos de Inteligência Artificial para o cálculo da probabilidade de Falha de um ativo depende do quão completo está o historial de falhas dos ativos do mesmo tipo. Se forem conhecidos parâmetros relevantes num momento próximo da falha de um ativo é possível formar um conjunto de amostras que contenha um conjunto de parâmetros e o correspondente estado de funcionamento ou de falha do ativo. Desta forma é possível construir um modelo de Classificação que tente extrair relações entre o valor desses parâmetros e a probabilidade de o ativo falhar. No âmbito do trabalho desenvolvido não foi possível obter um conjunto de dados que viabilizasse este tipo de estudo.

Os métodos de *Survival Analysis* têm uma vasta aplicação na análise de fiabilidade de equipamentos. A sua aplicação nos ativos de uma subestação elétrica permite estimar o seu tempo médio de vida e analisar a relação entre a sua idade e a probabilidade de terem uma falha geral que faça com que o ativo tenha que ser abatido. É possível aplicar estes métodos sabendo a idade em que um conjunto de ativos semelhantes falharam e a idade atual dos ativos em funcionamento.

Foi analisado um conjunto de disjuntores para aplicar métodos de *Survival Analysis* e comparar as suas previsões com a função de Sobrevivência definida empiricamente no algoritmo existente.

Não é possível comparar o tempo médio de vida calculado pelos métodos aplicados com as estimativas do algoritmo desenvolvido na EDP por não haver dados suficientes. O algoritmo utiliza vários parâmetros que não são conhecidos para os disjuntores abatidos.

4.2 Estimação da Probabilidade de Falha de um disjuntor através de *Survival Analysis*

Esta experiência visa estimar o efeito da idade de um disjuntor na sua probabilidade de falha. A figura 19 ilustra a contribuição da idade nos algoritmos existentes. De notar que a idade influencia diretamente a probabilidade de Falha e indiretamente através do Índice de Saúde.

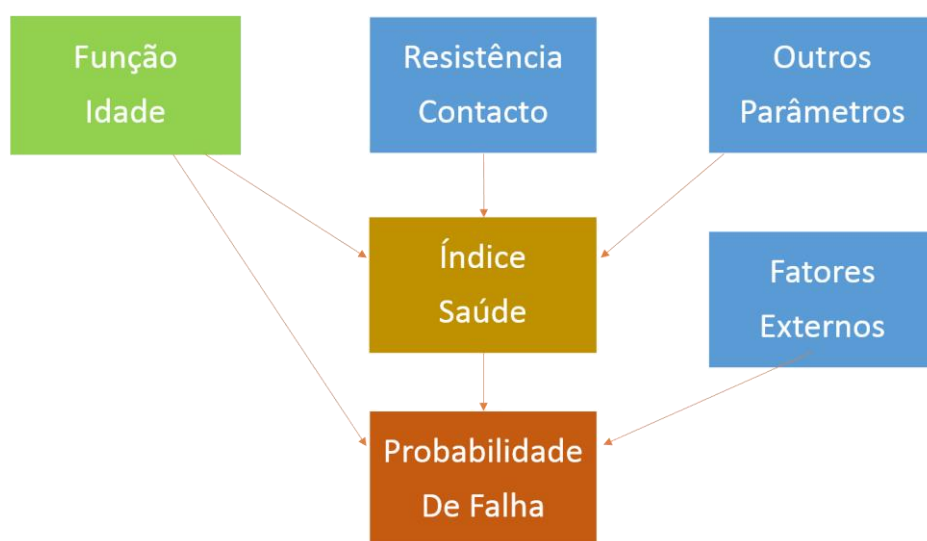


Figura 19 - Contribuição da idade do ativo nos algoritmos existentes

Como foi apresentado na secção 4.2, o algoritmo desenvolvido na EDP para estimação da probabilidade de falha de um ativo prevê um componente relativo à sua idade. A idade do ativo é convertida numa probabilidade de falha através de uma função exponencial, representada na figura 9. Esta falha é uma ocorrência que faz com que o ativo não cumpra com os padrões de funcionamento necessários, tendo que ser retirado.

A idade do ativo também é tida em conta no algoritmo relativo ao Índice de Saúde, no entanto, é necessário converter a probabilidade de falha dada pela função previamente referida num valor entre 0 e 100 em que quanto maior este for melhor é o estado do ativo.

Como o único evento considerado é a falha geral do ativo, qualquer que seja a sua causa, pode ser utilizado o complemento da função de probabilidade de falha para relacionar a probabilidade de sobrevivência com a idade, ou seja, se a probabilidade de um disjuntor com menos de dez anos falhar for de 10% então a probabilidade de esse disjuntor sobreviver até aos 10 anos é de 90%.

Esta função de Sobrevivência é equivalente às funções estimadas a partir de métodos de *Survival Analysis*. O objetivo deste estudo é implementar alguns destes métodos em dados reais e comparar a sua precisão com as do algoritmo existente.

4.2.1 Preparação dos dados

Para tornar possível esta experiência foi necessário reunir um conjunto de dados de sobrevivência sobre os disjuntores utilizados na rede elétrica da EDP Distribuição. Foi recolhida a informação dos cerca de 7000 disjuntores atualmente no ativo. Estes indivíduos são considerados censurados uma vez que a falha ainda não ocorreu, apenas se sabe a sua idade atual e, obviamente, que sobreviveram pelo menos até esse momento.

A existência de indivíduos não censurados é essencial para o sucesso deste tipo de estudo, no entanto, o sistema de informação utilizado pela EDP não guarda o histórico de ativos de uma subestação. Foi possível descobrir a data de abate de 31 disjuntores a partir de uma extração de formulários de abate. Devido à heterogeneidade das informações contidas em cada formulário não são conhecidos o meio de interrupção e o modelo de alguns destes disjuntores. São também tidos em conta 44 disjuntores previstos para abate em 2016 perfazendo um total de 75 disjuntores não censurados. Nos gráficos das figuras 20 e 21 são apresentadas as distribuições do meio de interrupção dos disjuntores abatidos e dos disjuntores em serviço.

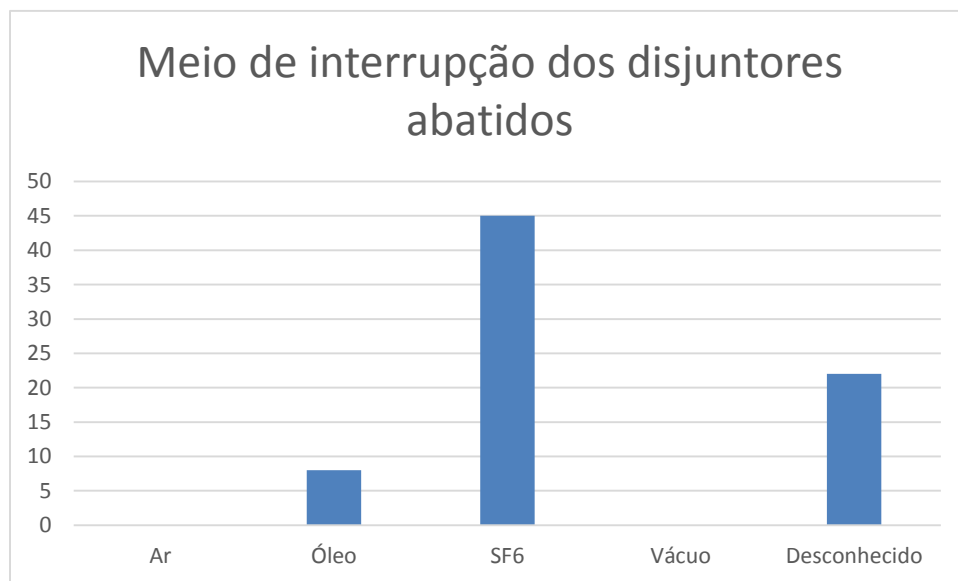


Figura 20 - Meio de interrupção dos disjuntores abatidos

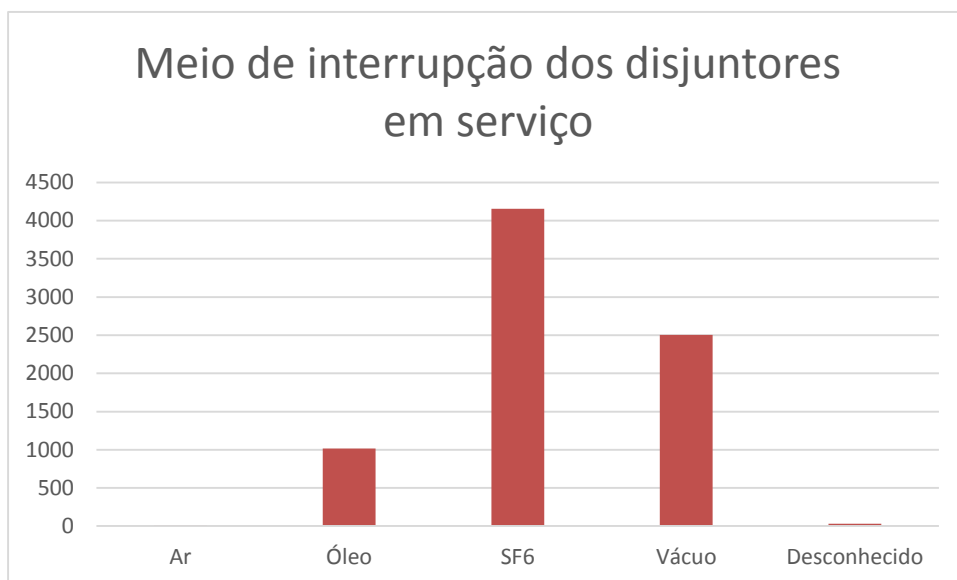


Figura 21 - Meio de interrupção dos disjuntores em serviço

Os dados recolhidos foram organizados numa tabela em que cada linha contém as informações de um disjuntor. Na tabela 16 é apresentada uma extração desta tabela. A variável Delta indica se o evento ocorreu no disjuntor.

<i>SAP ID</i>	<i>subestação</i>	<i>Data de entrada em serviço</i>	<i>Meio de interrupção</i>	<i>Ano de fabrico</i>	<i>Modelo</i>	<i>Data de Abate</i>	<i>Idade</i>	<i>Delta</i>
280202127	ARADA	14/06/1996	SF6	1995	FP1716B	23/11/2014	18	1
280199736	VÁRZEA	01/01/1996	SF6	1996	HA1/WS 17.06.16	24/11/2014	18	1
20141220045	VIDAGO	01/01/1994	SF6	1994	FP1716B	20/01/2015	21	1
20141218751	VIDAGO	01/01/1991	SF6	1991	FP1716B	20/01/2015	24	1
20141219040	AMARES	01/01/1989	SF6	1989	FP61B	20/01/2015	26	1
20141220239	GUIMARÃES	01/01/1987	SF6	1987	FP62D	20/01/2015	28	1
---	POMBAL	01/01/2005	---	2005	---	24/11/2014	9	1
---	POMBAL	01/01/1996	---	1996	---	25/11/2014	18	1
280218311	AGUEDA	30/12/2000	Vácuo	1998	Divac1716B	---	14	0
280202486	BARRÔ	30/09/1990	SF6	1988	FP61B	---	24	0
280819731	ALBERGARIA	21/12/2012	SF6	2012	SF1	---	2	0
280217656	VISTA ALEGRE	16/02/2007	SF6	2006	GL309-F1/2520	---	8	0
...

Tabela 16 - Dados de sobrevivência de disjuntores

4.2.2 Metodologia adotada

Para que os resultados dos modelos utilizados possam ser comparados de forma correta é necessário que estes sejam construídos e testados nas mesmas condições. Além disso, os dados de teste devem ser diferentes dos dados de treino para garantir que o modelo é capaz de generalizar as suas previsões para casos desconhecidos.

O método de validação utilizado foi o *10-fold Cross Validation*. Os dados são divididos em dez conjuntos, sendo os modelos treinados com nove dos conjuntos e o conjunto restante utilizado para validação. O processo é repetido para cada conjunto e o resultado final do modelo é a média de todas as iterações. Este método de validação é repetido dez vezes sendo também feita a média das estimativas para obter um valor único.

Para medir o erro dos modelos é utilizado o método *Brier Score* (Gerds and Schumacher 2007). Para um determinado momento t , é calculado o quadrado da diferença entre a previsão do modelo de o disjuntor sobreviver até a esse momento t e o estado observado do disjuntor em t , que pode ser 0 se o disjuntor já tiver falhado ou 1 se ainda estiver em funcionamento. Este valor é uma medida do erro logo quanto menor for melhor é a precisão do modelo. Para servir de referência, o *Brier Score* obtido por um método aleatório, como o lançamento de uma moeda, é igual a 33% e o de um modelo que preveja sempre uma probabilidade de 50% é de 25%. Um modelo ideal que acerte todas as previsões obtém um valor de 0%.

O *Brier Score* de cada modelo é calculado em vários momentos do ciclo de vida gerando-se assim uma curva do erro de predição. A área debaixo desta curva é igual ao erro cumulativo de predição, denominado de *Integrated Brier Score (IBS)*, que permite atribuir um valor único ao erro do modelo, representando o *Brier Score* do modelo ao longo do tempo de vida em estudo e facilitando a sua comparação com outros modelos.

Os métodos abordados foram os seguintes:

1. Kaplan-Meier;
2. Cox Proportional Hazards;
3. Algoritmo empírico, o método atualmente utilizado na EDP, igual ao complemento da função exponencial referida na secção anterior;
4. Distribuição Weibull;
5. Distribuição Exponencial;
6. Árvore de Decisão;
7. Random Survival Forest.

O método *Kaplan-Meier* e o algoritmo empírico são Não Paramétricos, apenas utilizam a idade do disjuntor para fazer a sua previsão. O método *Cox Proportional Hazards* e as distribuições de Weibull e Exponencial tiram partido do ano de fabrico dos disjuntores que é a única característica que foi possível identificar para todos os disjuntores. A Árvore de Decisão e

a *Random Survival Forest* são capazes de lidar com valores desconhecidos pelo que além do ano de abate também utilizam o modelo e o meio de interrupção dos disjuntores.

4.2.3 Resultados

Para comparar a capacidade preditiva dos modelos de *Survival Analysis* e do modelo empírico baseado no algoritmo existente foi utilizada a metodologia previamente apresentada, variando o número de indivíduos censurados tidos em conta. Foi definido um limite máximo de quarenta anos de idade devido a este ser o tempo de vida máximo de um disjuntor segundo os padrões da EDP Distribuição. Em suma, o procedimento é constituído pelos seguintes passos:

1. São escolhidos aleatoriamente n disjuntores em atividade, sendo a amostra total constituída por 75 disjuntores com data de abate conhecida juntamente com n indivíduos censurados;
2. Os modelos são validados através de *10-fold Cross Validation* utilizando o *Integrated Brier Score* como estimador do erro de previsão;
3. O passo 2 é repetido dez vezes e o valor final do erro de cada modelo é igual à média do seu *Integrated Brier Score* obtido em cada *10-fold Cross Validation* completa.

Os modelos foram testados com 0, 100, 2000 e 5000 indivíduos censurados. A tabela 17 apresenta o *Integrated Brier Score* obtido pelos modelos em cada caso e a evolução dos resultados é apresentada graficamente na figura 22.

Modelo	Número de indivíduos censurados			
	0	100	2000	5000
<i>Kaplan-Meier</i>	0.107	0.121	0.059	0.024
Algoritmo empírico	0.139	0.129	0.159	0.167
<i>Cox</i> com Ano de Fabrico	0.037	0.128	0.029	0.019
Árvore de Decisão	0.026	0.052	0.030	0.019
<i>Random Survival Forest</i>	0.033	0.042	0.013	0.008
<i>Weibull</i>	0.022	0.016	0.026	0.016
Exponential	0.105	0.142	0.053	0.025

Tabela 17 - Erro dos modelos por número de indivíduos censurados

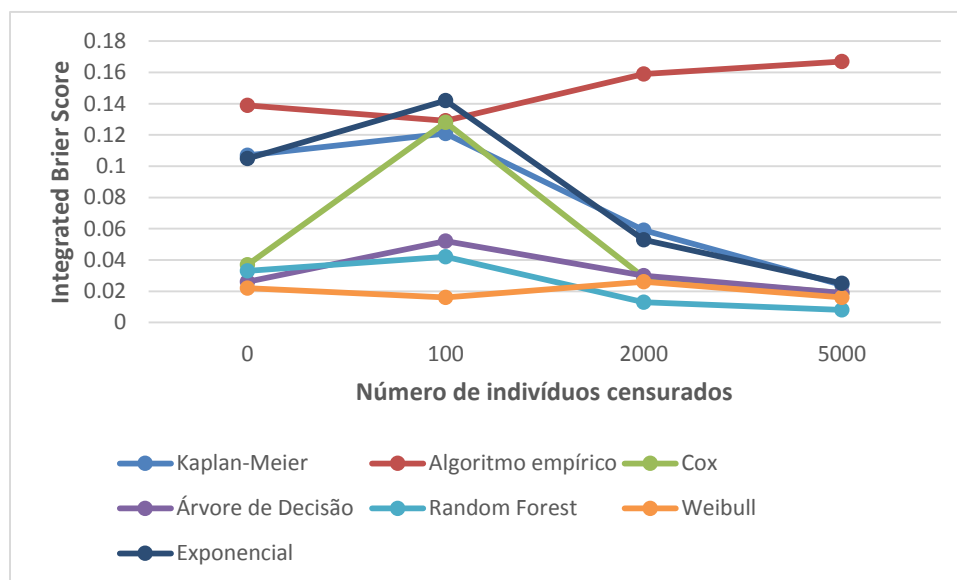


Figura 22 - Relação entre o IBS de cada modelo e o número de indivíduos censurados

Começando a análise pelo algoritmo empírico, o método desenvolvido pela EDP, este demonstra uma capacidade preditiva considerável, visto que o IBS não atinge valores elevados. Como seria de esperar, o erro não varia significativamente uma vez que as previsões são calculadas através de uma função definida previamente, pelo que este modelo não tem uma fase de treino propriamente dita, não tirando partido de um conjunto de treino maior.

Os modelos de *Survival Analysis* obtiveram um erro bastante menor. O indicador *Brier Skill Score* (BSS) permite analisar o erro de um modelo (BS) utilizando outro modelo como referência (BS_{ref}) comparando a sua capacidade preditiva. Este valor indica a precisão relativa de um modelo e pode tomar valores de $-\infty$ a 1, sendo que quanto maior o valor melhor é o modelo.

$$BSS = 1 - BS/BS_{ref} \quad (18)$$

A tabela 18 apresenta o *Brier Skill Score* de cada modelo com 5000 indivíduos censurados, com o algoritmo empírico como modelo de referência.

Modelo	<i>Brier Skill Score</i>
Algoritmo empírico	0
<i>Kaplan-Meier</i>	0.8563
<i>Cox com Ano de Fabrico</i>	0.8862
Árvore de Decisão	0.8862
<i>Random Survival Forest</i>	0.9521
<i>Weibull</i>	0.9042
Exponential	0.8503

Tabela 18 - *Brier Skill Score* dos modelos com 5000 indivíduos censurados

O modelo baseado em *Random Survival Forest* obteve os melhores resultados, tendo a distribuição de *Weibull* obtido também resultados muito bons. Os modelos *Kaplan-Meier*, *Cox* e a distribuição Exponencial tiveram um erro ligeiramente maior, verificando-se que com a redução do tamanho da amostra a diferença aumenta. A Árvore de Decisão não é tão afetada pelo tamanho da amostra mas não tem uma capacidade preditiva tão boa como a *Random Survival Forest* como seria de esperar uma vez que o intuito deste método é melhorar a capacidade de generalização dos modelos baseados em Árvores de Decisão.

É possível concluir que os modelos de *Survival Analysis* são efetivamente capazes de obter informação útil dos disjuntores ainda no ativo para melhorar a sua previsão em relação ao algoritmo empírico. Isto tem ainda maior relevância se se considerar o número reduzido de disjuntores abatidos em análise.

O processo de remodelação das Subestações elétricas geridas pela EDP Distribuição é uma iniciativa relativamente recente, pelo que é previsível que o número de disjuntores abatidos conhecidos aumente no futuro. De forma a perceber quais os métodos que mais têm a ganhar com este aumento foi realizada outra experiência. A metodologia de validação e de estimação do erro dos modelos foi a mesma, no entanto, em vez de variar o número de indivíduos censurados foram utilizados diferentes quantidades dos 75 disjuntores abatidos. A tabela 19 e a figura 23 apresentam os resultados obtidos pelos modelos com amostras constituídas por 200 indivíduos censurados e 25, 50 e 75 indivíduos não censurados.

Modelo	Número de indivíduos não censurados		
	25	50	75
<i>Kaplan-Meier</i>	0.099	0.117	0.123
Algoritmo empírico	0.144	0.133	0.131
<i>Cox</i> com Ano de Fabrico	0.241	0.176	0.145
Árvore de Decisão	0.076	0.034	0.041
<i>Random Survival Forest</i>	0.05	0.048	0.035
<i>Weibull</i>	0.016	0.016	0.017
Exponential	0.106	0.134	0.139

Tabela 19 - Erro dos modelos por número de indivíduos não censurados

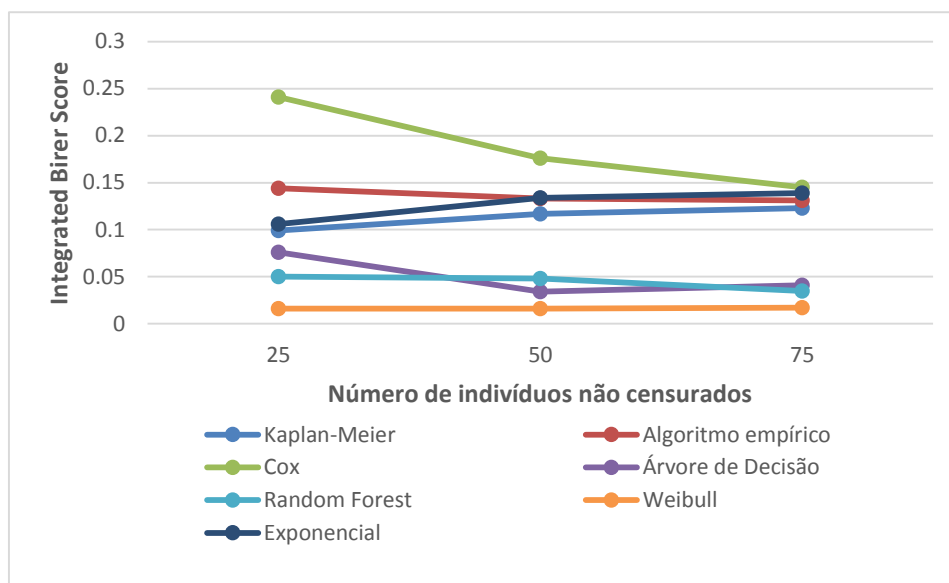


Figura 23 - Relação entre o IBS de cada modelo e o número de indivíduos não censurados

Nesta experiência verificam-se as mesmas diferenças entre os modelos. Os modelos com melhor resultado são a *Random Survival Forest*, a distribuição de *Weibull* e a *Árvore de Decisão*. De destacar que o erro da distribuição de *Weibull* se mantém constante enquanto a *Random Survival Forest* melhora a sua precisão com o aumento de indivíduos não censurados na amostra.

É preciso ter em conta que quanto maior o número de indivíduos não censurados mais fidedignos são os conjuntos de teste, uma vez que contêm mais casos com um tempo de vida efetivamente conhecido.

A distribuição de *Weibull* destaca-se pela sua capacidade preditiva, o que vem confirmar a sua adequação ao estudo do tempo de vida de uma população. O método baseado em *Random Survival Forest* também teve resultados excelentes. Seria interessante conhecer disjuntores não censurados com uma maior variedade de meios de interrupção pois poderia permitiria prever com mais precisão o tempo de vida por meio de interrupção.

Capítulo 5

Implementação do protótipo do Sistema de Apoio à Decisão

Neste capítulo é apresentada a implementação da solução para o problema abordado que consiste no desenvolvimento de um protótipo de um Sistema de Apoio à Decisão que automatize o cálculo do Índice de Saúde, da probabilidade de Falha e do tempo restante de vida dos ativos de uma subestação elétrica. Nas secções seguintes são apresentados os requisitos desta aplicação, as tecnologias utilizadas e as interfaces gráficas desenvolvidas.

5.1 Especificação de Requisitos

A EDP Distribuição é responsável por uma vasta rede de subestações, tendo que gerar uma grande quantidade de ativos complexos. Como já foi referido anteriormente, para fazer face às necessidades da empresa foram desenvolvidos algoritmos para caracterizar o estado desses ativos. Os dados de cada ativo são gerados a partir de várias fontes, como dados de cadastro, inspeções de rotina, ensaios e diagnósticos e monitorização *online*, sendo necessário centralizar esta informação e automatizar a aplicação dos algoritmos.

O protótipo desenvolvido tem que permitir não só a definição dos algoritmos a aplicar como também o cadastro das Subestações elétricas existentes e dos seus respetivos ativos. Além disso, a rede de subestações deve ser facilmente navegável, ajudando o utilizador a encontrar rapidamente o ativo que procura e a obter a informação de que necessita.

Para que a informação utilizada pelos algoritmos esteja acessível, é necessário permitir o registo dos valores medidos. Por exemplo, um Técnico deve poder aceder à aplicação e registar o resultado de um ensaio realizado num ativo.

Como a iniciativa de utilizar os indicadores referidos é relativamente recente, os algoritmos utilizados pela aplicação devem ser facilmente modificáveis de forma a refletirem a experiência adquirida com a utilização dos indicadores.

A correta implementação deste tipo de gestão permite identificar fatores de risco e priorizar as intervenções de manutenção de acordo com a condição dos ativos o que contribui para a redução dos custos, não só pelo decréscimo da ocorrência das falhas mas também pela diminuição de intervenções desnecessárias.

5.2 Tecnologias utilizadas

Tendo em conta os requisitos do sistema, optou-se por desenvolver uma aplicação *Web*. Desta forma, o protótipo pode ser utilizado em qualquer lugar, desde que haja ligação à Internet, permitindo aos vários intervenientes no processo de gestão acederem à informação quando têm necessidade de o fazer.

A aplicação foi desenvolvida em *Python* com recurso à *framework* Django. Esta *framework* acelera o processo de desenvolvimento fornecendo várias funcionalidades convenientes como a criação automática de tabelas na base de dados a partir das classes implementadas. Além disso, a linguagem *Python* tem uma grande popularidade na comunidade científica, pelo que existem várias bibliotecas que facilitam a análise estatística e a aplicação de métodos de *Data Mining* que possam vir a ser integrados no sistema.

As interfaces com o utilizador foram desenvolvidas com HTML e CSS e foi utilizado *Javascript* para melhorar a experiência do utilizador e facilitar a utilização da aplicação.

A base de dados foi implementada em *SQLite3* devido à sua elevada compatibilidade com o *Django*. Como esta aplicação é para ser utilizada dentro da EDP Distribuição o número de acessos concorrentes nunca será elevado pelo que esta base de dados se afigura como a melhor opção por não necessitar de instalação nem configuração.

De forma a garantir a correta implementação dos algoritmos foram implementados testes unitários para comparar os resultados obtidos com os resultados calculados manualmente em casos de teste.

5.3 Modelo de dados

O modelo de dados da aplicação pode ser dividido em duas partes. Numa são modelados os algoritmos, sendo definidos os tipos de ativo e os parâmetros necessários para caracterizar a sua condição. Por outro lado, é necessário permitir a criação das Subestações elétricas e dos seus

respetivos ativos. Na figura 24 é apresentado o diagrama de classes utilizado para definir o algoritmo correspondente a cada tipo de ativo.

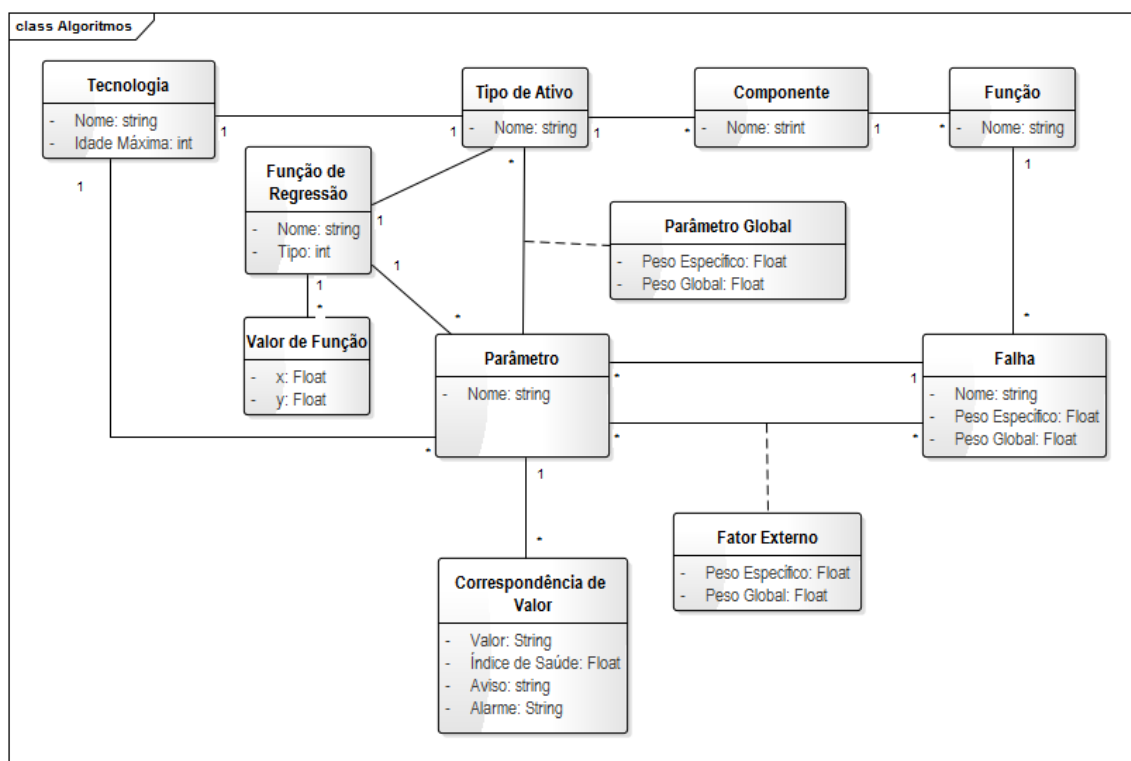


Figura 24 - Diagrama de classes para modelação dos algoritmos

A classe “Tipo de Ativo” representa um dos ativos de uma subestação. Como foi referido na secção 3.2, um ativo é decomposto em componentes que por sua vez podem ter uma ou mais funções com os seus respectivos modos de falha. Estas entidades são representadas através das classes “Componente”, “Função” e “Falha”.

Um tipo de ativo pode ainda ter várias tecnologias, como no caso do disjuntor onde são previstas as tecnologias SF6, Vácuo e Óleo, pelo que é necessário diferenciar estes ativos. As tecnologias são associadas através da Classe “Tecnologia” onde também é definido o seu tempo de vida máximo.

Na base de tudo estão os parâmetros, representados na classe “Parâmetro”. Para permitir a atribuição de Índices de Saúde aos estados ou intervalos de valores que um parâmetro pode tomar é utilizada a classe “Correspondência de Valor”. É possível, também, atribuir uma mensagem de alarme ou de aviso a um dado valor.

Caso o Índice de Saúde de um parâmetro seja calculado através de uma função esta é criada através da classe “Função de Regressão”. São escolhidos pontos que relacionem os valores numéricos que o parâmetro pode tomar com o respetivo Índice de Saúde, por exemplo 0 anos de idade têm um Índice de Saúde máximo. Estes pontos são criados através da classe “Valor de

Função”. A regressão utilizada para estimar o Índice de Saúde a partir dos pontos definidos é identificado através do atributo “Tipo” da tabela “Função de Regressão”. Foram implementados os seguintes tipos de regressão: Linear, Exponencial, Quadrática de 2º Grau e Quadrática de 3º Grau.

Pode haver quatro tipos de parâmetros:

- Parâmetros que influenciam o Índice de Saúde de uma falha;
- Fatores externos, utilizados para estimar a probabilidade de uma falha ocorrer. São representados através da classe “Fator Externo” onde é atribuído o seu peso específico e global.
- Parâmetros de envelhecimento, que são associados a uma tecnologia e representam os fatores que influenciam o tempo de vida útil do ativo;
- Parâmetros Globais, que são associados diretamente a um tipo de ativo através da classe “Parâmetro Global” e afetam o seu Índice de Saúde global e o específico de cada componente.

Na figura 25 são apresentadas as classes utilizadas para representar o conjunto de ativos a avaliar.

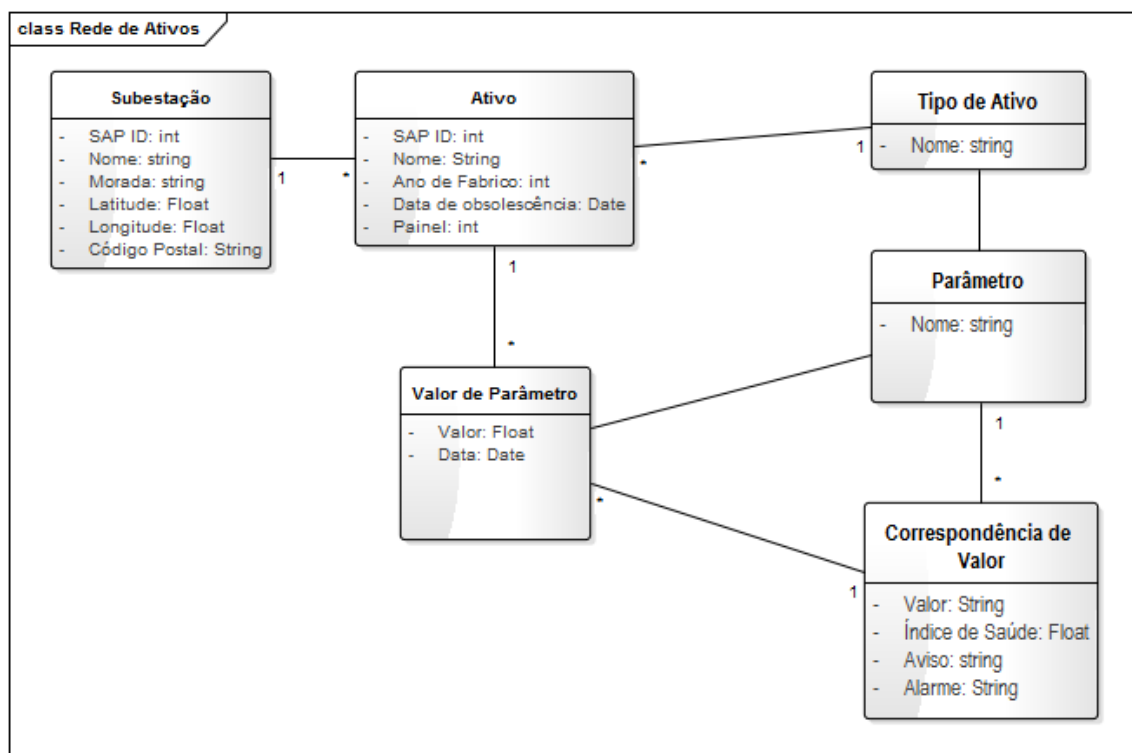


Figura 25 - Diagrama de classes da rede de ativos

A classe “subestação” representa uma subestação elétrica, sendo guardados o seu SAP ID, que é o seu número identificativo na Rede de Subestações, o seu nome, a morada e posição geográfica. Os ativos de uma subestação são criados através da classe “Ativo”, onde são atribuídos o nome, o ano de fabrico e a data em que se prevê que este se torne obsoleto. Caso se aplique é também atribuído um SAP ID e o número do painel em que está inserido na subestação.

Cada ativo está associado a um tipo de ativo e os valores registados para os parâmetros correspondentes são guardados através da classe “Valor de Parâmetro”. Se o valor do Parâmetro for um intervalo ou um estado então é atribuído o Índice de Saúde correspondente na classe “Correspondência de Valor”. Caso seja uma função então é utilizado o campo “Valor” e o Índice de Saúde é obtido através da função atribuída ao parâmetro correspondente.

Através desta arquitetura é possível estender e modificar de forma simples os algoritmos de cada tipo de ativo. Os pesos atribuídos a cada parâmetro e os seus valores correspondentes podem ser facilmente alterados. É possível acompanhar o aparecimento de novas tecnologias ou até mesmo de novos tipos de ativo sem que seja necessário mudar o código fonte, sendo apenas necessário introduzir novos algoritmos na base de dados.

5.4 Interfaces gráficas

As interfaces implementados podem ser agrupadas em dois conjuntos: o Portal, onde são apresentadas as Subestações e os seus ativos e o Painel de Administração, que oferece as funcionalidades de definição dos algoritmos e de inserção de dados.

5.4.1 Portal

O Portal é constituído por três páginas: a página inicial, a página de subestação e a página de Ativo. Através destas páginas um utilizar pode navegar de forma intuitiva e prática pela rede de subestações e consultar o estado dos ativos. Nesta secção serão apresentadas estas páginas e a informação que estas contêm.

5.4.1.1 Página inicial

Na página inicial da aplicação, apresentada na figura 26, é possível navegar pela rede de subestações carregadas no sistema. Uma subestação pode ser acedida de duas maneiras, navegando pelo mapa e carregando no marcador correspondente ou utilizando a listagem no painel “Subestações”. O campo de pesquisa neste painel permite pesquisar as subestações pelo nome ou pelo seu SAP ID como se pode verificar na figura 27.

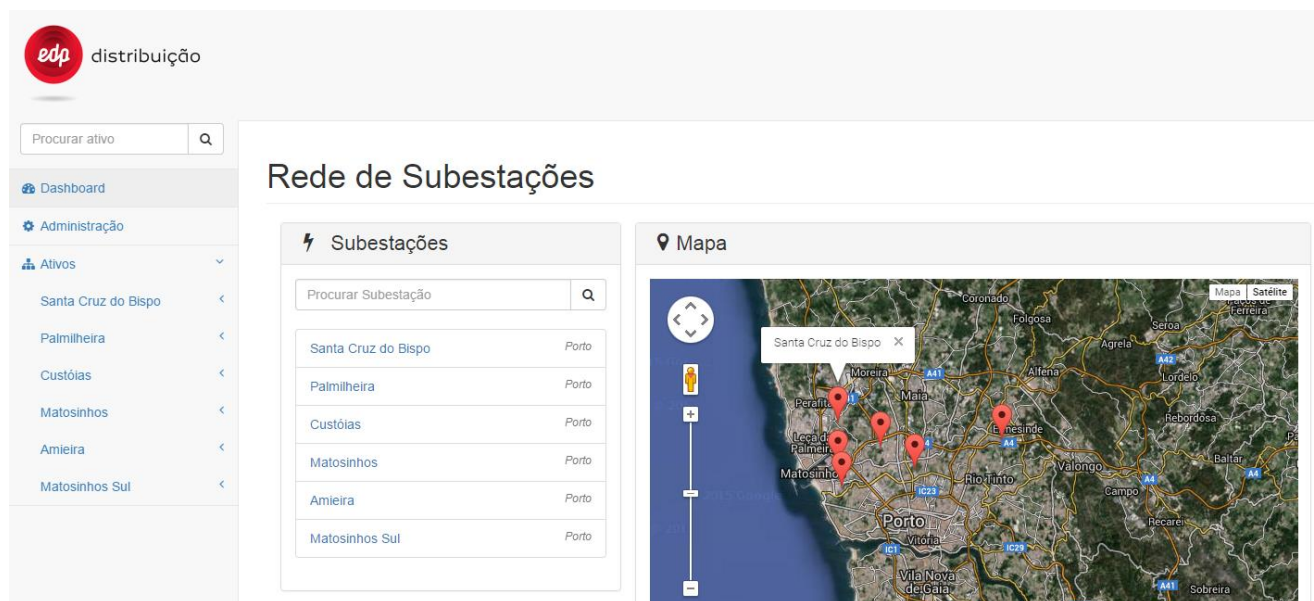


Figura 26 - Página Inicial

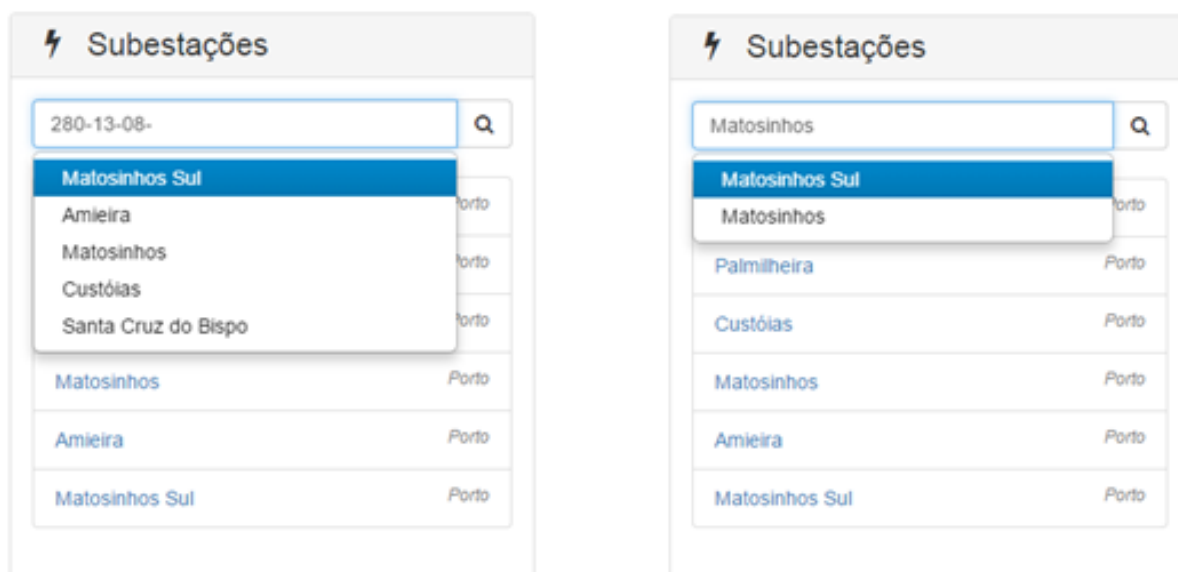


Figura 27 - Pesquisa de Subestação por nome ou por SAP ID



SAP ID	Tipo	Ativo	Índice de Saúde	Probabilidade de Falha	Tempo de Vida
280202124	Disjuntor - Óleo	Disjuntor Óleo	43	78%	5
Teste2	Bateria - Alcalina	Bateria Alcalina Teste	65	49%	2
280202126	Disjuntor - Vácuo	Disjuntor Vácuo	70	34%	18

Figura 28 - Página Inicial - Ativos em risco

Na página inicial são listados os ativos com maior probabilidade de falhar para que o gestor possa identificar de forma rápida os maiores pontos de risco na rede como se pode ver na tabela da figura 28.

A barra lateral está presente em todas as páginas do portal e permite ao utilizador visitar o painel de administração através do botão “Administração”. Caso se deseje visitar um ativo específico é possível pesquisar diretamente pelo seu nome ou *id* através da caixa de texto no topo da barra lateral. Neste menu é também possível percorrer a “árvore” de ativos, abrindo a aba da subestação em que o ativo está colocado como é demonstrado na figura 29.

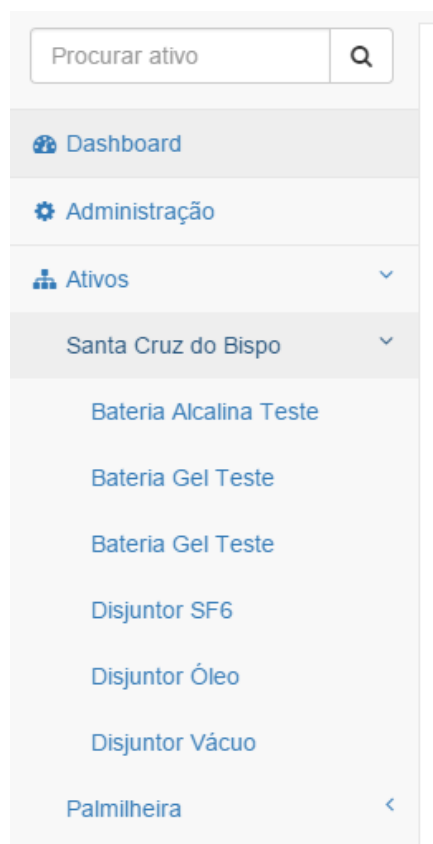


Figura 29 - Pesquisa em árvore dos ativos

5.4.1.2 Página de subestação

q

Subestação Santa Cruz do Bispo

Informações [Editar](#)

SAP ID: 280-13-08-S-5-0059
Morada: Rua do Azemel, s/nº
Código Postal: 4455-976
GA: Porto

Ativos

SAP ID	Tipo	Ativo	Índice de Saúde	Probabilidade de Falha	Tempo de Vida
Teste	Bateria - Gel	Bateria Gel Teste	90	19%	4
280202125	Disjuntor - SF6	Disjuntor SF6	82	24%	23
280202126	Disjuntor - Vácuo	Disjuntor Vácuo	70	34%	18
Teste2	Bateria - Alcalina	Bateria Alcalina Teste	65	49%	2
280202124	Disjuntor - Óleo	Disjuntor Óleo	43	78%	5

Mapa

Figura 30 - Página de Subestação

Na página de uma subestação são apresentadas as suas informações, como o seu SAP ID, a morada e a sua localização da subestação no mapa, podendo este ser navegado para visitar as subestações circundantes. Além disso, são listados os ativos presentes na subestação sendo possível verificar rapidamente o seu Índice de Saúde, Probabilidade de Falha e tempo de vida restante. Para visitar a página de um dos ativos basta carregar na linha correspondente da tabela.

O botão “Editar” no painel “Informações” direciona o utilizador para a página da subestação no Painel de Administração, sendo possível modificar as suas informações.

5.4.1.3 Página do Ativo

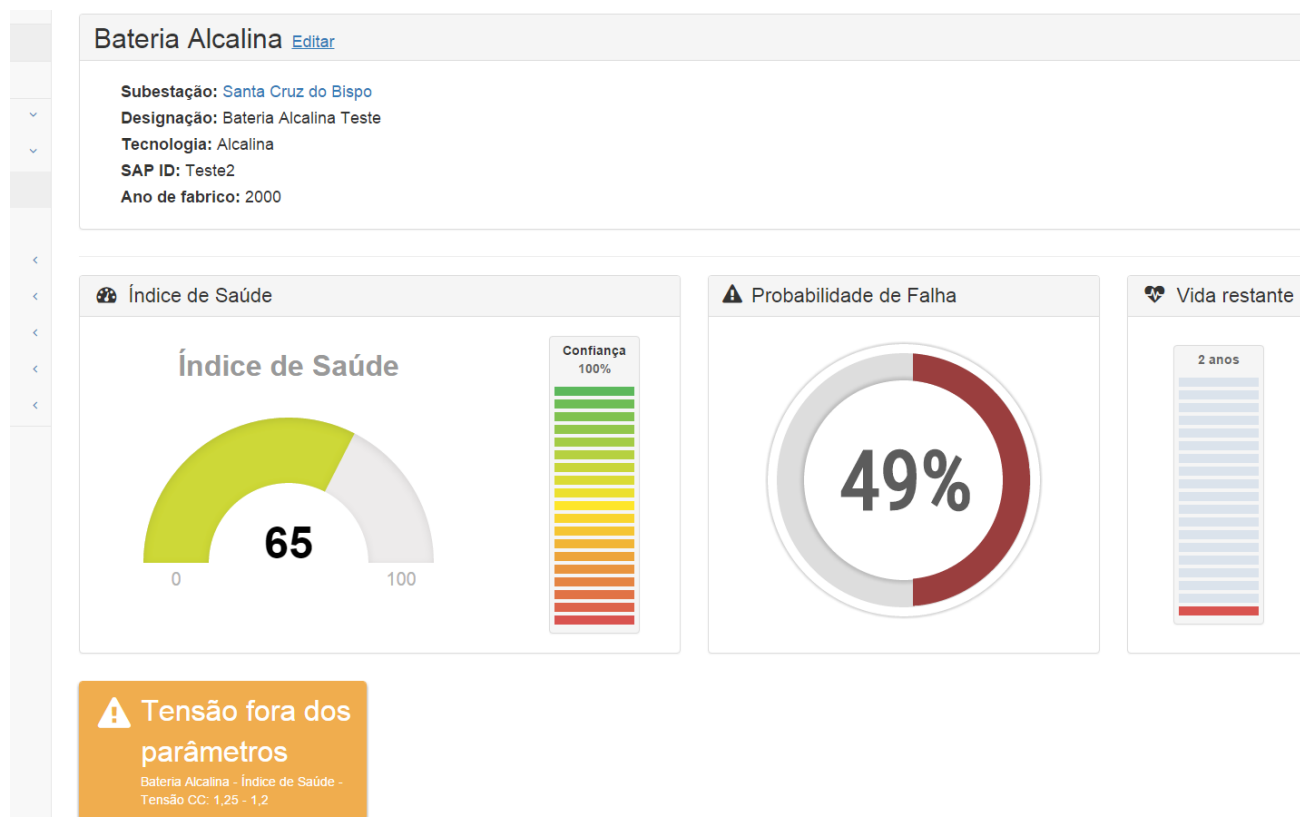


Figura 31 - Página de Ativo - Indicadores principais

A figura 31 mostra o topo da página de um ativo. São apresentadas as informações disponíveis como a sua designação, o seu ano de fabrico ou a subestação a que o ativo pertence. O botão “Editar” permite alterar as informações do ativo e registar valores medidos para os seus parâmetros.

Os principais indicadores são representados graficamente de forma intuitiva. O Índice de Saúde é apresentado em conjunto com o nível de confiança, que indica o quão fiável é o indicador.

O preenchimento da barra no painel “Vida Restante” é relativa ao tempo de vida máximo definido para a tecnologia do ativo. No caso da imagem apresentada o ativo é uma Bateria Alcalina, cujo tempo de vida máximo definido no algoritmo é de 20 anos. A estimativa do tempo de vida restante do ativo em questão é de 2 anos, portanto, a barra está preenchida a 10%.

Depois dos principais indicadores, são apresentados os alarmes e alertas despoletados pelos parâmetros atuais do ativo.

Na figura 32 é apresentado o segundo conjunto de painéis da página de um ativo. Os indicadores são apresentados de forma mais detalhada para dar uma visão mais completa sobre o ativo ao utilizador.



5.4.2 Painel de Administração

No Painel de Administração é possível definir os algoritmos e introduzir os dados relativos às Subestações elétricas e aos seus principais ativos.

Nesta secção serão apresentadas as páginas mais relevantes.

5.4.2.1 Página Inicial

Home	
Site administration	
Algoritmos	
Falhas	+ Add ≡ Change
Parâmetros	+ Add ≡ Change
Tecnologias	+ Add ≡ Change
Tipos de Ativo	+ Add ≡ Change
Authentication and Authorization	
Groups	+ Add ≡ Change
Users	+ Add ≡ Change
Rede de Ativos	
Ativos	+ Add ≡ Change
Subestações	+ Add ≡ Change
Utils	
Funções de Regressão	+ Add ≡ Change
Recent Actions	
My Actions	
≡ Bateria - Alcalina	Tipo de Ativo
≡ Bateria - Gel	Tipo de Ativo
≡ Alimentador - Tipo	Tipo de Ativo
≡ Alimentador - Não Tipo	Tipo de Ativo
≡ Disjuntor - SF6	Tipo de Ativo
≡ Disjuntor - Vácuo	Tipo de Ativo
≡ Disjuntor - Óleo	Tipo de Ativo
≡ Bateria Alcalina Teste	Ativo
≡ Bateria Gel Teste	Ativo
≡ Disjuntor	Tipo de Ativo

Figura 34 - Página Inicial do Painel de Administração

Como se pode verificar na figura 34, o Painel de Administração está dividido em quatro componentes:

- “Algoritmos”, que permite a criação de tipos de ativos e os parâmetros a ser utilizados para estimar os indicadores referidos;
- “Authentication and Authorization”, onde são atribuídos acessos a utilizadores para que estes possam utilizar as funções do Painel de Administração;
- “Rede de Ativos”, onde são criadas as subestações e os respetivos ativos, permitindo atribuir valores aos parâmetros utilizados pelos algoritmos desenvolvidos.
- “Utils”, que possibilita a criação das funções a ser utilizadas pelos algoritmos.

Nesta página são também apresentadas as últimas ações efetuadas, como a modificação de algoritmos ou a atribuição de um valor a um parâmetro de um ativo.

5.4.2.2 Página de Tipo de Ativo

Home > Algoritmos > Tipos de Ativo > Bateria - Alcalina

Change Tipo de Ativo History

Nome: Bateria - Alcalina

Tecnologia: Alcalina

Função de Envelhecimento: Bateria Alcalina Envelhecimento

Parâmetro	Peso Específico	Peso Global
Bateria - Alcalina - Índice de Saúde - Vida	20.0	20.0

Add another Parâmetro Global

Componentes: Componente Baterias

Add another Componente

Figura 35 - Página de Tipo de Ativo

Nesta página é possível definir um tipo de ativo, incluindo todas os componentes necessários para a aplicação dos algoritmos desenvolvidos pela EDP Distribuição. Na figura 35 é apresentada a página correspondente ao tipo de ativo “Bateria Alcalina”. Nesta página é atribuído a tecnologia ao ativo, neste caso “Alcalina”, e a função de envelhecimento. Na aba seguinte são definidos os parâmetros globais que afetam o Índice de Saúde do ativo, incluindo o seu peso específico e global.

Na aba “Components”, que pode ser visualizada aberta na figura 365, é feita a decomposição do ativo em componentes, funções e falhas. Para cada falha é atribuído um peso local e global, que vão ser utilizados para o cálculo do Índice de Saúde, e os pesos a atribuir à idade do ativo e ao Índice de Saúde no cálculo da probabilidade de essa falha ocorrer.

Na figura 37 é demonstrado o formulário de criação dos parâmetros que afetam cada modo de falha. São previstos dois modos de valorização dos parâmetros:

- Atribuição de uma função, através do campo “Function”;
- Definição de intervalos de valores e o Índice de Saúde correspondente.

Os níveis de aviso e de alarme são também aqui definidos.

Componentes

Componente Baterias

Componente Baterias

Funções

Função Assegurar a alimentação DC

Nome Assegurar a alimentação DC

Falhas

Falha Autonomia insuficiente

Nome Autonomia insuficiente

Peso Global 40.0

Peso Específico 40.0

Peso Condição 55.0

Peso Idade 20.0

Parâmetros

Parâmetro Bateria - Alcalina - Índice de Saúde - Tempo de descarga

Parâmetro

Add another Parâmetro

Falha Tensão fora dos parâmetros

Falha

Add another Falha

Figura 36 - Página de Tipo de Ativo - Definição dos componentes

Parâmetros

Parâmetro Bateria - Alcalina - Índice de Saúde - Tempo de descarga

Nome Tempo de descarga

Função de Regressão

Correspondências de Valor

Valor	Índice de Saúde	Aviso	Alarme
20	100.0		
15	75.0	Autonomia insuficiente	
10	50.0	Autonomia insuficiente	
5	25.0		Autonomia insuficiente
0	0.0		Autonomia insuficiente

Add another Correspondência de Valor

Parâmetro

Add another Parâmetro

Figura 37 - Página de Tipo de Ativo - Definição de Parâmetro

5.4.2.3 Página de criação de subestação


Home > Network > Substations > Santa Cruz do Bispo

Change substation History

Instalação: Santa Cruz do Bispo

Coordenadas: Latitude: 41.21205 Longitude: -8.6775000000000

Start typing an address ...



Rua Azemel 35, 4455 Santa Cruz do Bispo, Portugal

SAP ID: 280-13-08-S-5-0059

Morada: Rua do Azemel, s/nº

C.P.: 4455-976

GA: Porto

Delete Save and continue editing Save and add another Save

Figura 38 - Página de criação de Subestação

Através da página apresentada na figura 38, é possível definir as informações de uma subestação elétrica. É suportada a atribuição de coordenadas geográficas para localizar mais fidedignamente a subestação no mapa.

5.4.2.4 Página de criação de Ativo

Um ativo pertencente a uma subestação é definido preenchendo o formulário apresentado na figura 39. Além das informações que caracterizam o ativo, como o seu SAP ID ou o ano de fabrico, é atribuído um dos tipos de ativo criados anteriormente. Desta forma são-lhe associados o conjunto de parâmetros a que devem ser atribuídos valores para caracterizar o seu estado.

A atribuição dos valores é feita através da metade inferior do formulário. Como se pode verificar nas figuras 40 e 41, apenas são disponibilizados para escolha os parâmetros e valores correspondentes ao tipo de ativo. Caso o parâmetro seja valorizado através de uma função o seu preenchimento é feito através do campo da coluna “Value”.

5.4.2.5 Página de criação de funções

Home > Utils > Funções de Regressão > Bateria Alcalina Envelhecimento

Change Função de Regressão History

Tipo de Regressão: Quadrática 3º Grau

Nome: Bateria Alcalina Envelhecimento

Valores de Função		
X	Y	
0.0	5.0	X
5.0	20.0	X
10.0	40.0	X
15.0	70.0	X
20.0	100.0	X
		—
		—
		—

[Add another Valor de Função](#)

Delete
Save and continue editing
Save as new
Save

Figura 42 - Página de criação de funções

As funções utilizadas nos algoritmos são criadas a partir da página apresentada na figura 42. A relação entre duas variáveis é definida através de um conjunto de pontos, sendo escolhido o tipo de regressão a utilizar. O exemplo apresentado representa a definição da função de envelhecimento de uma bateria alcalina, em que no eixo dos “x” está a idade do ativo e no eixo dos “y” a probabilidade deste falhar.

Capítulo 6

Conclusões e Trabalho futuro

6.1 Conclusões

Para o desenvolvimento do trabalho realizado, foi fundamental o estudo aprofundado do contexto da EDP Distribuição e dos processos de gestão em uso. Foi implementado um protótipo de um Sistema de Apoio à Decisão com o objetivo de auxiliar os processos de tomada de decisões para a gestão dos principais ativos de uma subestação elétrica.

A aplicação desenvolvida caracteriza de forma fundamentada e objetiva a condição dos ativos referidos através da estimação de três indicadores:

1. Índice de Saúde;
2. Probabilidade de Falha;
3. Tempo de vida restante.

A sua estimação é feita através da aplicação de algoritmos desenvolvidos pela empresa. Através de uma análise estatística de ensaios em disjuntores foi possível contribuir para a parametrização do seu algoritmo de estimação do Índice de Saúde.

Quando corretamente especificados, os indicadores referidos são os alicerces para um modelo de manutenção preditiva, com o foco em atuar antes que as falhas ocorram e evitar custos desnecessários.

As principais vantagens da aplicação do protótipo no contexto da EDP são:

1. Rapidez de acesso à condição de um ativo;
2. Facilidade de modificação dos algoritmos para refletir a experiência adquirida neste tipo de manutenção;
3. Possibilidade de inserção direta dos parâmetros necessários, o que pode ser efetuado pelo Técnico que recolhe a informação;
4. Centralização da informação das Subestações existentes e dos seus respetivos ativos;

5. Redução do tempo gasto no processo de decisão na gestão dos ativos, já que reduz os intervenientes necessários e facilita a recolha e análise dos dados;
6. Facilidade de acompanhamento da evolução dos indicadores, permitindo identificar processos de degradação.

Foi feita a revisão de abordagens na área de *Data Mining* para melhoria da estimação dos indicadores referidos. A aplicação dos métodos identificados no contexto da EDP foi estudada e foi concebida uma experiência para comparar a estimação da probabilidade de um disjuntor ter que ser abatido consoante a sua idade, tendo sido obtidos melhores resultados do que os conseguidos pelo algoritmo empírico existente.

6.2 Desenvolvimentos futuros

O próximo grande passo a dar seria a implementação de um sistema informático mais abrangente, capaz de adquirir de forma automática os parâmetros necessários para o cálculo dos indicadores referidos. Isto exigiria um processo de informatização dos dados de ensaios existentes e a implementação de novas tecnologias de monitorização, como a utilização de sensores capazes de registar os parâmetros relativos a cada ativo. A aplicação implementada facilita esta tarefa uma vez que permite o registo dos parâmetros numa base de dados, sendo necessário implementar uma API que faça a interface as fontes de informação e a base de dados.

Outra adição interessante na aplicação desenvolvida seria a inclusão do cálculo do risco de um ativo através da avaliação da relação entre a probabilidade de este falhar e as consequências dessa falha.

Quanto a abordagens de *Data Mining* para a estimação dos indicadores, os métodos de Aprendizagem Supervisionada mostram-se promissores, especialmente para a previsão da probabilidade de falha e do tempo de vida restante. No entanto, é necessário um processo abrangente de recolha de dados que permita a obtenção de resultados significativos. A estimação do Índice de Saúde apresenta mais dificuldades devido à maior complexidade deste indicador.

Referências

- 60599, Standard IEC. 2007. "Guide for the interpretation of dissolved gas analysis and gas-free."
- Anders, G., S. Otal, and T. Hjartarson. 2006. "Deriving asset probabilities of failure: effect of condition and maintenance levels." Power Engineering Society General Meeting, 2006. IEEE, 0-0 0.
- Arshad, M., S. Islam, and A. Khaliq. 2014. "Fuzzy logic approach in power transformers management and decision making." *Dielectrics and Electrical Insulation, IEEE Transactions on* 21 (5):2343-2354. doi: 10.1109/TDEI.2014.003859.
- Ashkezari, A. D., Ma Hui, T. K. Saha, and C. Ekanayake. 2013. "Application of fuzzy support vector machine for determining the health index of the insulation system of in-service power transformers." *Dielectrics and Electrical Insulation, IEEE Transactions on* 20 (3):965-973. doi: 10.1109/TDEI.2013.6518966.
- Ashkezari, A. D., Ma Hui, T. K. Saha, and Cui Yi. 2014. "Investigation of feature selection techniques for improving efficiency of power transformer condition assessment." *Dielectrics and Electrical Insulation, IEEE Transactions on* 21 (2):836-844. doi: 10.1109/TDEI.2013.004090.
- Bacha, Khmais, Seifeddine Souahlia, and Moncef Gossa. 2012. "Power transformer fault diagnosis based on dissolved gas analysis by support vector machine." *Electric Power Systems Research* 83 (1):73-79. doi: <http://dx.doi.org/10.1016/j.epsr.2011.09.012>.
- Badune, J., S. Vitolina, and V. Maskalonoks. 2013. "Methods for Predicting Remaining Service Life of Power Transformers and Their Components." *Power and Electrical Engineering* 31:123-126.
- Breiman, Leo. 2001. "Random Forests." *Mach. Learn.* 45 (1):5-32. doi: 10.1023/a:1010933404324.
- Cota-Felix, E. Jorge E., Franciso Rivas-Davalos, Serguei Maximov. 2009. "A new method to evaluate mean life of power system equipment". *CIREN, 20th International Conference on Electricity Distribution* (1):8-11
- Dan, Zhou, Wang Zhongdong, and Li Chengrong. 2013. "Data Requisites for Transformer Statistical Lifetime Modelling—Part I: Aging-Related Failures." *Power Delivery, IEEE Transactions on* 28 (3):1750-1757. doi: 10.1109/TPWRD.2013.2264143.
- Dorison, Eric, Frederic Lesur, Dominique Meurice and Gao Roinel. 2007. "Health Index". Jicable 07
- Fayyad, Usama, Gregory Piatetsky-shapiro and Padhraic Smyth. 1996. "From Data Mining to Knowledge Discovery in Databases". *AI Magazine* 17 (1):37-54

- Ganyun, L. V., Cheng Haozhong, Zhai Haibao, and Dong Lixin. 2005. "Fault diagnosis of power transformer based on multi-layer SVM classifier." *Electric Power Systems Research* 74 (1):1-7. doi: <http://dx.doi.org/10.1016/j.epsr.2004.07.008>.
- Gerds, Thomas A., and Martin Schumacher. 2007. "Efron-Type Measures of Prediction Error for Survival Analysis." *Biometrics* 63 (4):1283-1287. doi: 10.1111/j.1541-0420.2007.00832.x.
- Gulati, R., and R. Smith. 2009. *Maintenance and Reliability Best Practices*: Industrial Press.
- Han, Han, Wang Hou-jun, and Dong xiucheng. 2011. "Transformer Fault Dignosis Based on Feature Selection and Parameter Optimization." *ICSGCE 2011*:27-30.
- Han, Jiawei and Micheline Kamber. 2006. "Data Mining: Concepts and Techniques", Second Edition, 2006
- Hhleln, I., A. J. Kachler, S. Tenbohlen, and M. Stach. 2003. "Transformer Management German Experience with Condition Assessment." *Contribution for CIGRE SC 12/A2*.
- Hjartarson, T., and S. Otal. 2006. "Predicting Future Asset Condition Based on Current Health Index and Maintenance Level." *Transmission & Distribution Construction, Operation and Live-Line Maintenance, 2006. ESMO 2006. IEEE 11th International Conference on*, 15-19 Oct. 2006.
- Insights, Asset. 2003. "Predictive Maintenance." Asset Insights Accessed 16 Jun.
- Ishwaran, Hemant, Udaya B. Kogalur, Eugene H. Blackstone, and Michael S. Lauer. 2008. "Random survival forests." *The Annals of Applied Statistics* 2 (3):841-860.
- Jahromi, A., R. Piercy, S. Cress, J. Service, and W. Fan. 2009. "An approach to power transformer asset management using health index." *Electrical Insulation Magazine, IEEE* 25 (2):20-34. doi: 10.1109/MEI.2009.4802595.
- Kaplan, E. L., and P. Meier. 1958. "Nonparametric estimation from incomplete observations". *J. Amer. Statist. Assn.* 53 (282): 457-481
- Kleinbaum, David G., and Mitchel Klein. 2012. "Survival Analysis A self learning tex"t, Third edition, XV, 700p
- LaCommare, Kristina Hamachi, and Joseph H. Eto. 2004. "Understanding the Cost of Power Interruptions to U.S. Electricity Consumers " *Ernest Orlando Lawrence, Berkeley National Laboratory, Environmental Energy Technologies Division*.
- Lawless, J.F.. 1982." Statistical Models and Methods for Lifetime Data". Wiley, New York.
- Lee, Seung-Hyuk, In-Su Bae, Jin-O Kim, and Sung-Hoon Lee. 2008. "Evaluation aging failure probability of generating units using data analytic method." *European Transactions on Electrical Power* 19 (5):631-641.
- Lee, Tsair-fwu, Ming-yuan Cho, Fu-min Fang, and Chang Gung. 2007. "Features Selection of SVM and ANN Using Particle Swarm Optimization for Power Transformers Incipient Fault Symptom Diagnosis." *International Journal of Computational Intelligence Research* 3 (1):60-65.

- Lian, Chao, and Lin Ma. 2013. "Health assessment model of power transformer based on dissolved gas analysis by support vector machine." *Information Management, Innovation Management and Industrial Engineering (ICIII)*, 2013 6th International Conference on, 23-24 Nov. 2013.
- Liu, Jianqing, Kai Zheng, Hao Zhang, and Daogang Peng. 2008. "A Comparative Research on Power Transformer Fault Diagnosis Based on Several Artificial Neural Networks." *Journal of Computational Information Systems* 9 (18):7501-7508.
- Maciejewski, H., G. J. Anders, and J. Endrenyi. 2011. "On the use of statistical methods and models for predicting the end of life of electric power equipment." *Power Engineering, Energy and Electrical Drives (POWERENG)*, 2011 International Conference on, 11-13 May 2011.
- Mukti, PH, FA Pamuji, and BS Munir. 2014. "Implementation of Artificial Neural Networks for Determining Power Transformer Condition." *ADCONIP 2014*.
- Nemeth, B., S. Laboncz, I. Kiss, and G. Csepes. 2010. "Transformer condition analyzing expert system using fuzzy neural system." *Electrical Insulation (ISEI)*, Conference Record of the 2010 IEEE International Symposium on, 6-9 June 2010.
- Qing, Li, Yang Wei, Zhu Guang-xun, and Zhou Jie. 2011. "Study on survival analysis method for equipment life and its application." *Industrial Engineering and Engineering Management (IE&EM)*, 2011 IEEE 18Th International Conference on, 3-5 Sept. 2011.
- Reclamation, Bureau of. 2003. "Transformer Diagnostics." *Facilities, Instructions, Standards and Technique (FIST)* 3 (31).
- XiaoDong, Yu, and Zang Hongzhi. 2008. "Transformer fault diagnosis based on rough sets theory and artificial neural networks." *Condition Monitoring and Diagnosis*, 2008. CMD 2008. International Conference on, 21-24 April 2008.
- XiaoDong, Yu, and Zhang Li. 2009. "Transformer Fault Diagnosis Based on Improved SVM Model." *Natural Computation*, 2009. ICNC '09. Fifth International Conference on, 14-16 Aug. 2009.

Anexo A

Análise de ensaios em disjuntores

8.1 Tempo de Abertura

8.1.1 Meio de interrupção

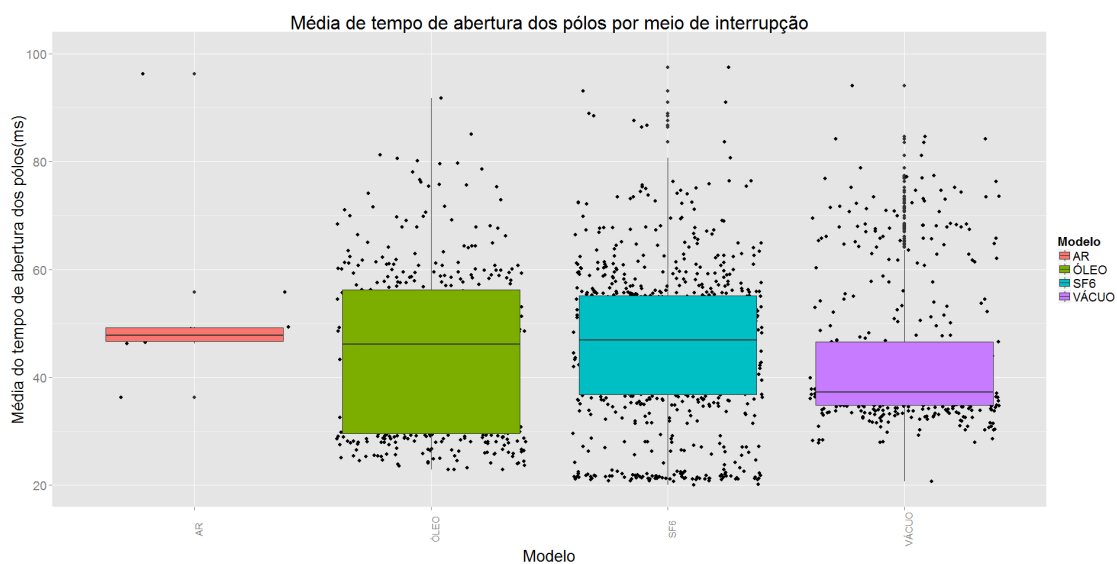


Figura 43 - Tempo de Abertura por meio de interrupção

8.1.2 Média Tensão

8.1.2.1 Vácuo

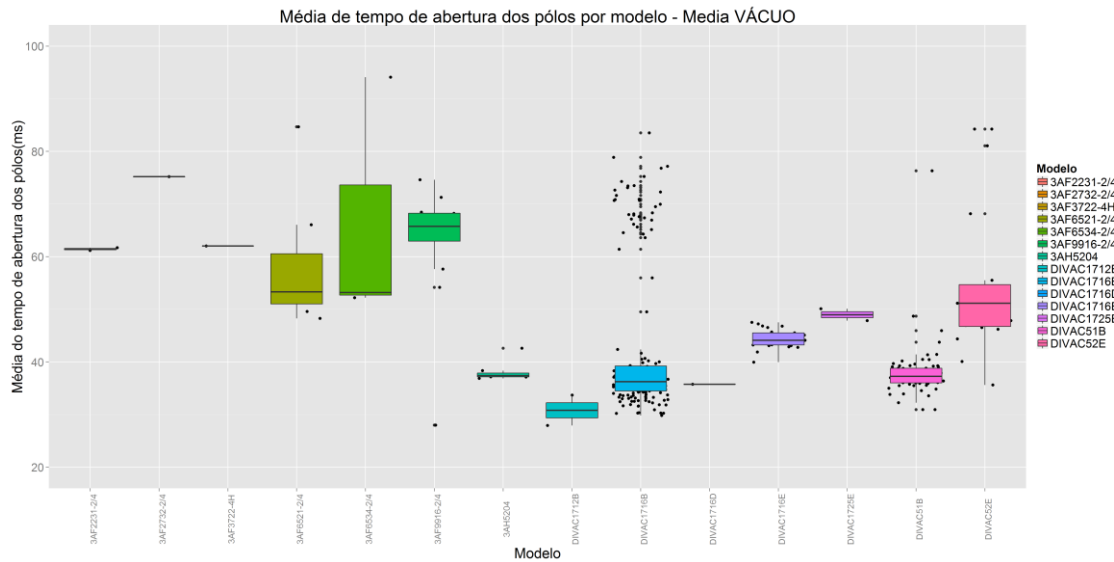


Figura 44 - Tempo de Abertura de disjuntores a Vácuo de Média Tensão

8.1.2.2 SF6

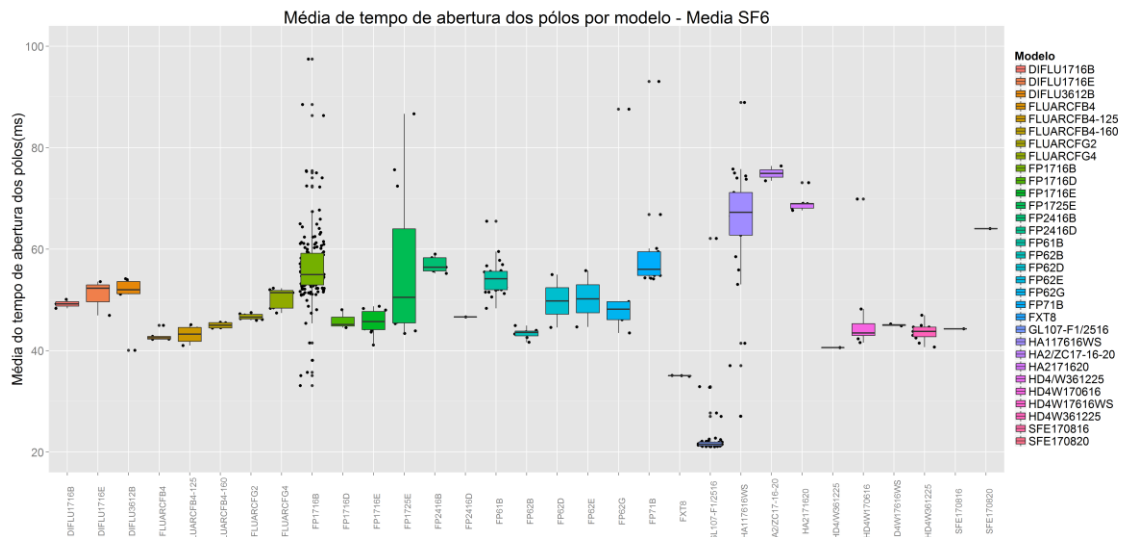


Figura 45 - Tempo de Abertura de disjuntores a SF6 de Média Tensão

8.1.2.3 Óleo

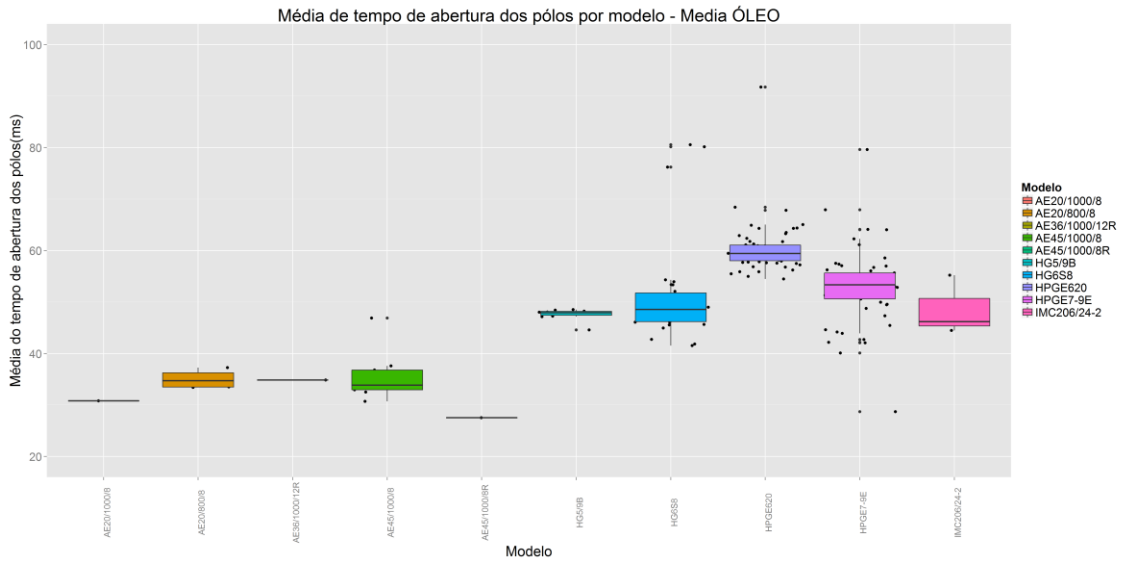


Figura 46 - Tempo de Abertura de disjuntores a Óleo de Média Tensão

8.1.3 Alta Tensão

8.1.3.1 SF6

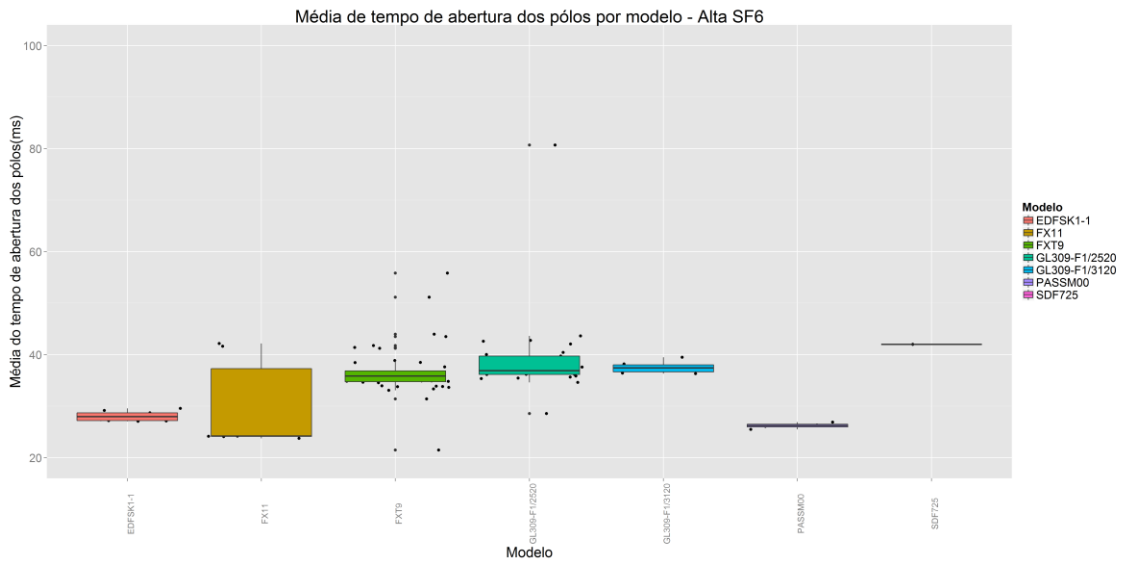


Figura 47 - Tempo de Abertura de disjuntores a SF6 de Alta Tensão

8.1.3.2 Óleo

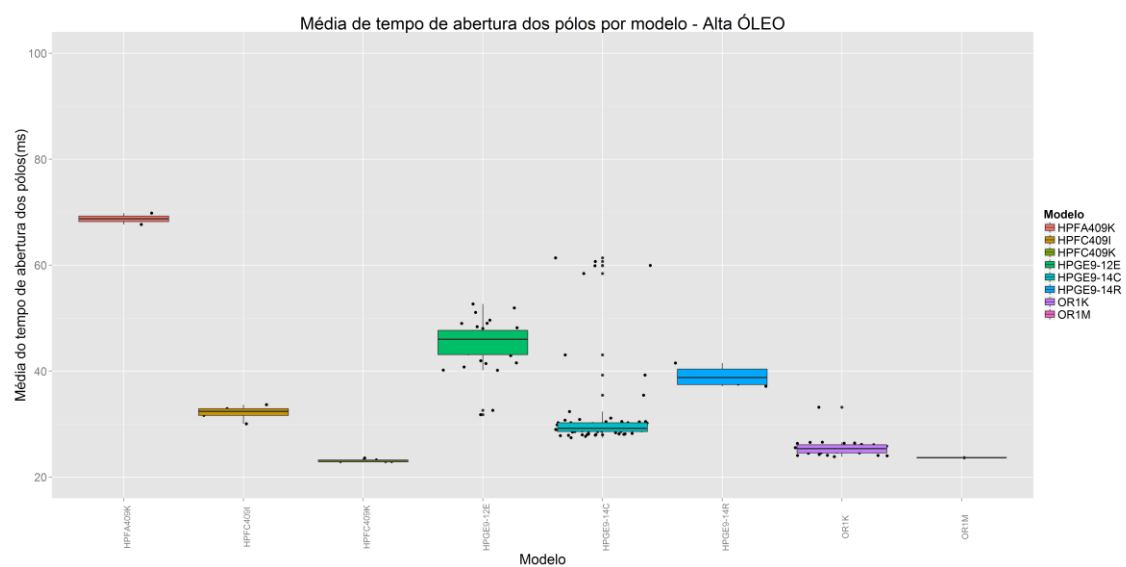


Figura 48 - Tempo de Abertura de disjuntores a Óleo de Alta Tensão

8.2.1.3 Óleo

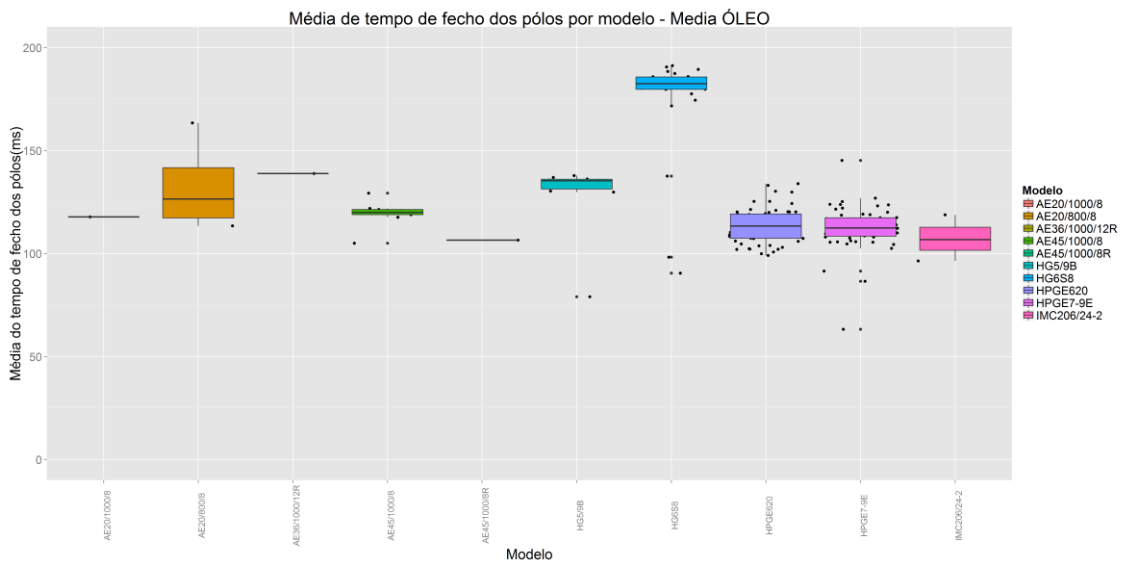


Figura 51 - Tempo de Fecho de disjuntores a Óleo de Média Tensão

8.2.2 Alta Tensão

8.2.2.1 SF6

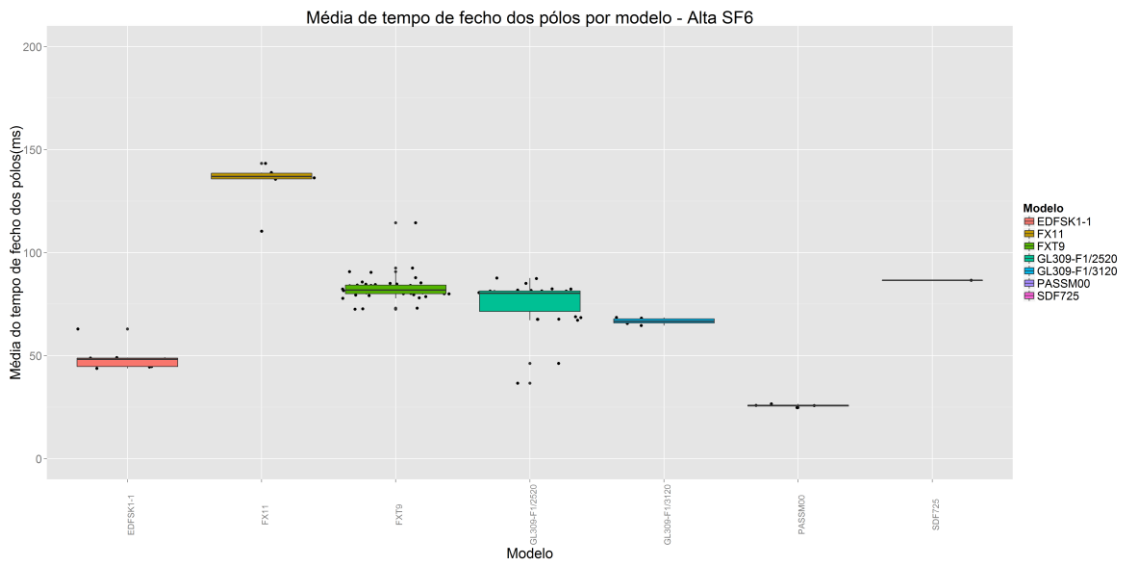


Figura 52 - Tempo de Fecho de disjuntores a SF6 de Alta Tensão

8.2.2.2 Óleo

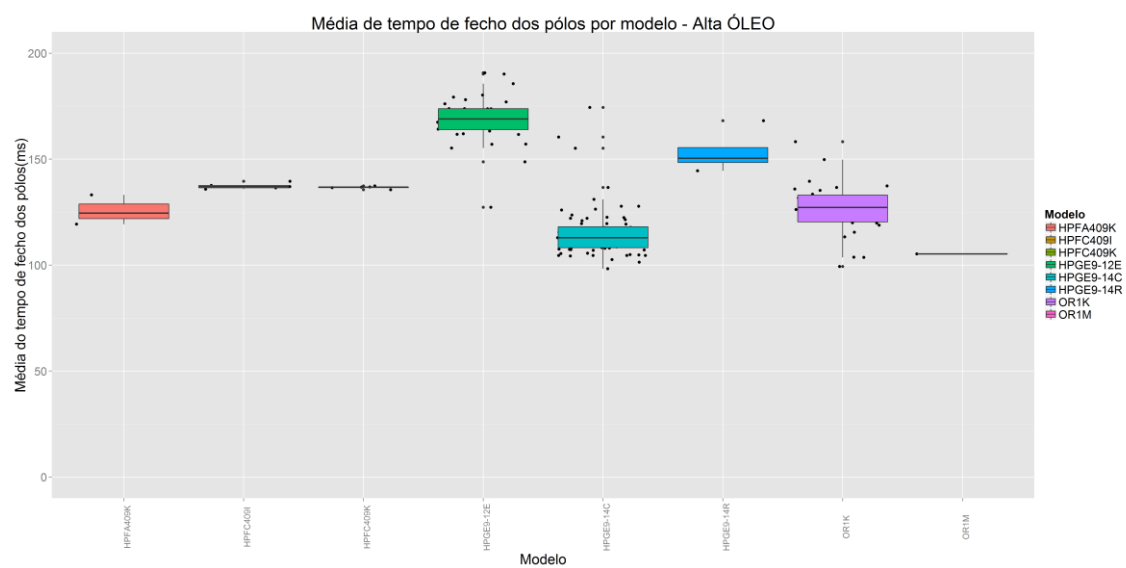


Figura 53 - Tempo de Fecho de disjuntores a Óleo de Alta Tensão

8.3.1.3 Óleo

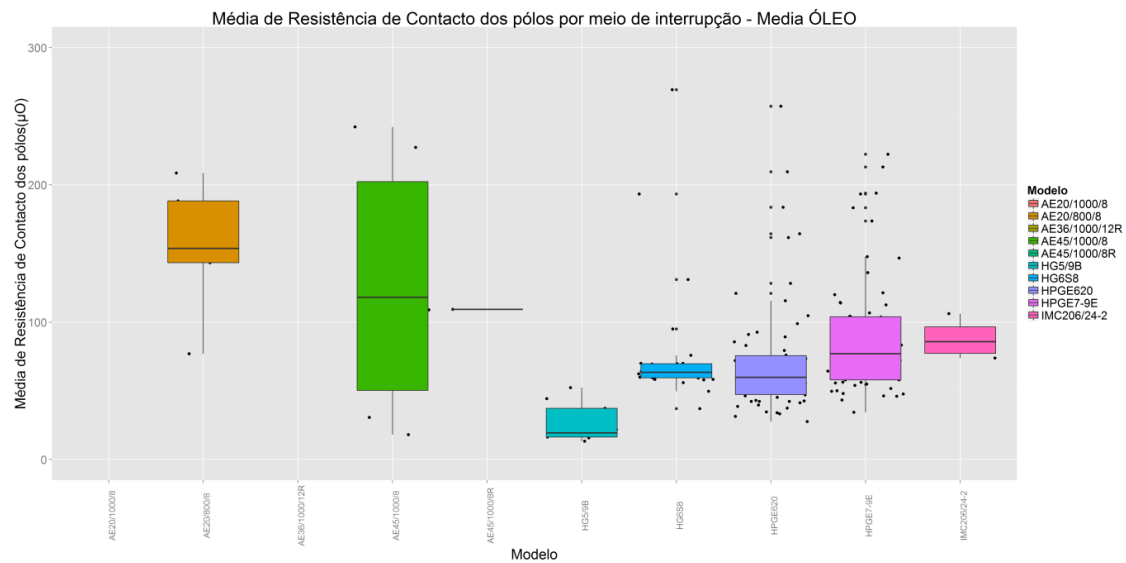


Figura 56 - Resistência de Contacto de disjuntores a Óleo de Média Tensão

8.3.2 Alta Tensão

8.3.2.1 SF6

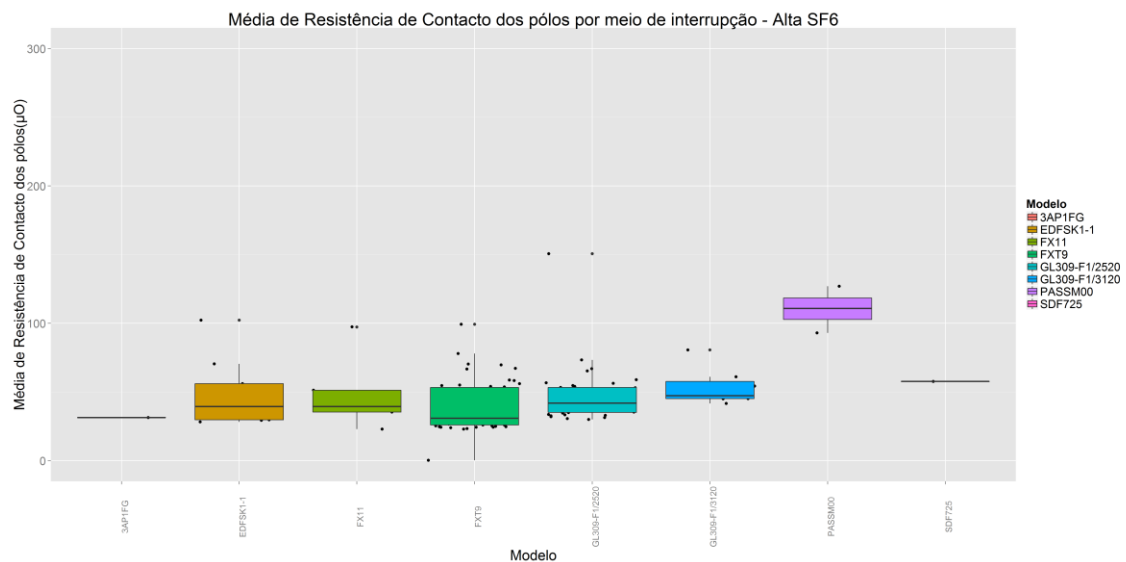


Figura 57 - Resistência de Contacto de disjuntores a SF6 de Alta Tensão

8.3.2.2 Óleo

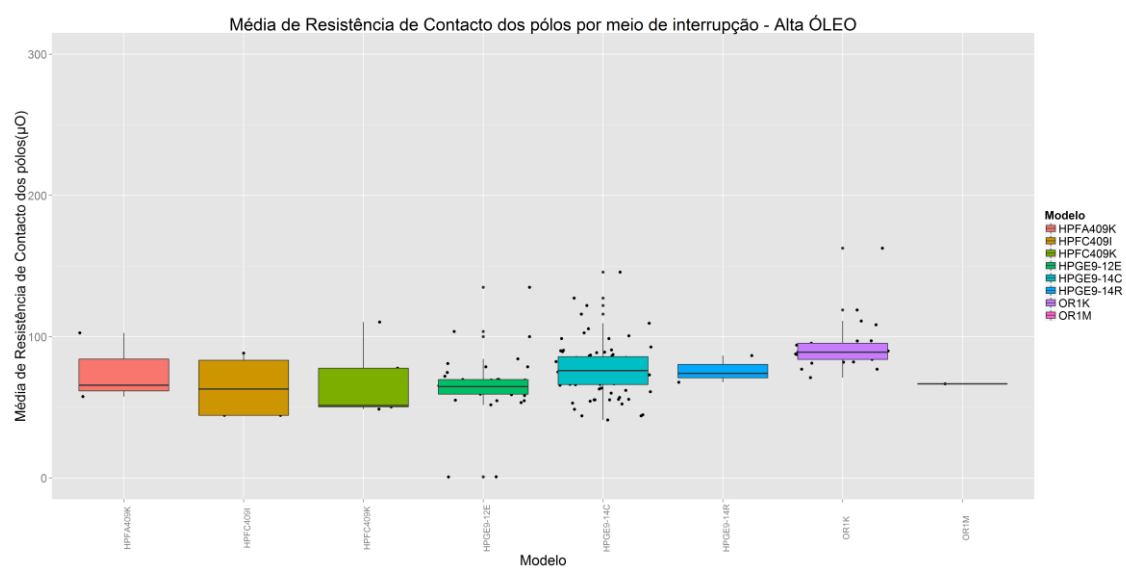


Figura 58 - Resistência de Contacto de disjuntores a Óleo de Alta Tensão